

The evolution of determinate and indeterminate nodules within the Papilionoideae subfamily

Guiling Ren

Thesis committee

Promotor

Prof. Dr T. Bisseling
Professor of Molecular Biology
Wageningen University & Research

Co-promotor

Dr R. Geurts
Associate professor, Laboratory of Molecular Biology
Wageningen University & Research

Other members

Prof. Dr M.E. Schranz, Wageningen University & Research
Prof. Dr J.I. Sprent (Emeritus), University of Dundee, Scotland UK
Prof. Dr S. Goormachtig, VIB / Ghent University, Belgium
Dr E.K. James, the James Hutton Institute, Dundee, Scotland UK

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

The evolution of determinate and indeterminate nodules within the Papilionoideae subfamily

Guiling Ren

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 23 January 2018
at 4.00 p.m. in the Aula.

Guiling Ren

The evolution of determinate and indeterminate nodules within the Papilionoideae subfamily, 130 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2018)

With references, with summary in English

ISBN: 978-94-6343-230-6

DOI: 10.18174/429101

CONTENTS

Outline	7
Chapter 1	
General Introduction	9
Chapter 2	
A technology platform for <i>Indigofera argentea</i>	23
Chapter 3	
Variation in infectiveness of rhizobial strains nodulating the promiscuous legume <i>Indigofera argentea</i>	47
Chapter 4	
Evolution of determinate and indeterminate nodules in Papilionoideae	63
Chapter 5	
The role of NCRs in bacteroid differentiation in <i>Indigofera argentea</i> root nodules	87
Chapter 6	
General Discussion	105
Acknowledgements	125
<i>Curriculum vitae</i>	127
Education statement	128

Outline

We selected the legume *Indigofera argentea* for our research. It was collected in the desert of Jizan in Saudi Arabia. We selected this species for two reasons. It is well adapted to heat and drought and therefore it has the potential to study, in the future, mechanisms that confer tolerance to these abiotic stresses. Further, *Indigofera* represents an early branching lineage within the indigoferoid/milletioid clade. Therefore, *Indigofera* is a key genus in studying the evolution of nodulation within the Papilionoideae subfamily.

In Chapter 1, a general introduction is given on nitrogen fixing symbiosis of legumes and rhizobia. In this introduction, we focus on the process of nodule initiation and organogenesis. Two main nodule types, determinate and indeterminate nodule, are introduced based on the knowledge of the few well-studied legumes species. Further, terminal differentiation of rhizobia that is induced by NCR peptides of the host is introduced.

In Chapter 2, we characterized the desert legume *I. argentea* and developed a platform by which future studies on mechanisms controlling abiotic stress become available. We developed an *Agrobacterium rhizogenes*-mediated root transformation procedure and did a *de novo* transcriptome assembly using RNA of various organs.

In Chapter 3, about 60 rhizobium strains have been isolated from nodulated *I. argentea* plants that were collected in the desert. The strains were characterized by 16S sequencing and their nodulation abilities were studied. One of the efficient nodulating *Bradyrhizobium* strains was selected for further studies.

In Chapter 4, nodule development of *Indigofera* and *Tephrosia* species was analyzed and described in detail. Species from both genera have an indeterminate growth. However, it was shown that this is not due to a meristem that is formed at the primordium stage, which is the basis of indeterminate growth of IRLC species like *Medicago*. The indeterminate growth was shown to be due to secondary clusters of dividing infected cells that were formed from nodule parenchyma cells. Therefore, it evolved independently from the indeterminate growth from IRLC species which is controlled by a persistent meristem composed of non-infected cells.

In Chapter 5, it is shown that bacteroids from *I. argentea* nodules are markedly enlarged as described for IRLC (e.g. *Medicago*) and *Aeschynomene* species. This increase in size was correlated with endoreduplication and terminal differentiation. Further, it was shown that in nodules of *I. argentea* 4 NCR genes are expressed. Therefore it is probable that these terminal differentiation is controlled by the NCR peptides and this evolved independently in the IRLC clade, *Aeschynomene* and *Indigofera*.

In Chapter 6, I discuss my results and put them in a broader perspective. I summarize and discuss the determinate and indeterminate nodule evolution in the Papilionoideae subfamily.

We have shown in Chapter 4 that *Indigofera* and *Tephrosia* species have a nodule development that is very similar to that of determinate nodules and their indeterminate growth is due to the formation of secondary clusters of dividing cells. Species from other clades of the Papilionoideae subfamily were analyzed. Based on these analyses, I conclude that the ancestor of the Papilionoideae subfamily formed determinate nodules that had the ability to form secondary clusters of dividing cells. Further, NCR-triggered bacteroid evolution in the Papilionoideae subfamily is discussed.

Chapter 1

General Introduction

1. Biological nitrogen fixation

Atmospheric dinitrogen gas (N_2) can be reduced to ammonia (NH_3) by several bacteria. This process is called biological nitrogen fixation. An efficient process of biological nitrogen fixation occurs in nodule symbiosis of bacteria and plants. Nitrogen-fixing symbiosis between bacteria and plants include for example, actinorhizal nodule formation by members of the actinobacterial genus *Frankia*, and legume nodule formation with rhizobia. Biological nitrogen fixation is catalyzed by an enzyme complex called nitrogenase. The subunits of this enzyme are encoded by *nif* genes which are conserved in nitrogen-fixing microorganisms (Normand and Bousquet, 1989). Biological nitrogen fixation is of importance in agriculture and can reduce the use of mineral nitrogen (e.g. NH_4^+ , NO_3^-) (Fields, 2004; Novotny et al., 2010).

The rhizobium-legume interaction is the best studied and agronomically most important nitrogen-fixing symbiosis. Several α - and β -proteobacteria species, collectively named rhizobia, can establish a symbiosis with legumes. The rhizobia induce the formation of nodules on roots of the legume plants and in specialized nodule cells they fix nitrogen into NH_4^+ , which is assimilated by the host plant. In return, the plant provides the required carbon sources among others for the high energy demanding nitrogen-fixation process.

2. Legume nodules

The Leguminosae family is the third-largest land plant family containing more than 770 genera and 19,500 species (Lewis et al., 2005, 2013; LPWG, 2017). The majority of them can be nodulated by rhizobia. The Leguminosae family is divided into six subfamilies: Caesalpinioideae, Cercidoideae, Detarioideae, Dialioideae, Duparquetioideae, and Papilionoideae (LPWG, 2017; Sprent et al., 2017). The Papilionoideae, the largest subfamily, is divided into distinct clades; the genistoids, dalbergioids, indigoferoids, milletioids, robinoids, and IRLC (Fig. 1). The genistoids consists of several genera, among which *Lupinus* has been widely studied for its nodulation characteristics. In the dalbergioids, *Aeschynomene* is the best studied genus. The indigoferoids, containing the genus *Indigofera*, represents an early branching lineage within the indigoferoid/milletioid clade (Wojciechowski et al., 2004; Lavin et al., 2005). The milletioids contains important crops like soybean and common bean. The robinoids contains for example *Sesbania*, *Robinia*, and the model legume *Lotus*. The IRLC clade contains the model legume *Medicago* and crops like pea.

Nodulation of legumes by rhizobia involves two main processes, namely, bacterial infection and nodule formation. Legume nodules are generally grouped into two main types, namely determinate and indeterminate, based on whether a persistent meristem is formed. These

different nodule types will be described in detail in paragraph 3. Here, I will first discuss the infection process and the early stages of nodule/primordium formation.

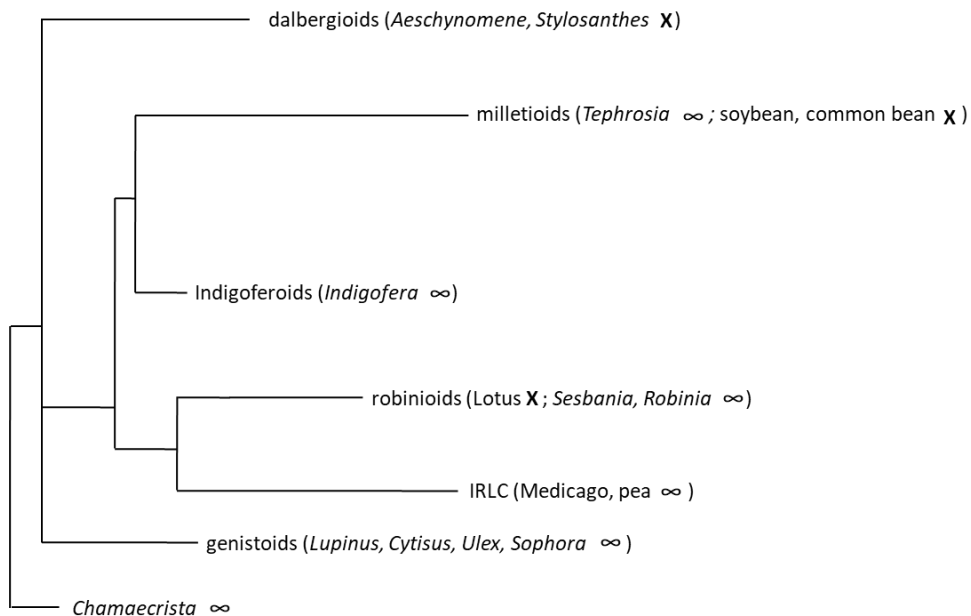


Fig. 1 The Papilionoideae subfamily is divided into distinct groups. Species/genera belonging to the groups are indicated (between brackets). Whether indeterminate (∞) or determinate (x) nodules are formed by the species/genera in these groups is indicated next to the genus name. The schematic representation of the Papilionoideae subfamily is drawn based on a phylogenetic tree from Oono et al. (Oono et al., 2010).

Rhizobial infection, for example of the model legumes (e.g. Medicago, Lotus, soybean), is initiated at the sites where a deformed root hair forms a curl (Callaham and Torrey, 1981). Subsequently, infection threads (ITs) are formed which are surrounded by a plant-derived membrane and cell wall material (Rae et al., 1992). The ITs are filled with dividing bacteria and they grow to the base of the root hair containing epidermal cell (Brewin, 2004; Gage, 2004). In some legumes, rhizobia enter their host by other mechanisms, such as transcellular penetration (crack-entry) between cells. This is for example the case in some *Sesbania*, *Aeschynomene* and *Lupinus* species (Gonzalez-Sama et al., 2004; Bonaldi et al., 2010; Capoen et al., 2010).

The ITs or penetrating bacteria traverse inwards to the root cortex where cortical cell division is initiated which results in a nodule primordium. The primordia can be induced in different cortical cell layers in different species. In soybean, the nodule primordia are formed in cortical cells directly adjacent to the infected root hair cells and subsequently differentiate

into a determinate nodule. In *Lotus*, nodule primordia are formed from the middle cortex and also develop into a determinate nodule. *Sesbania* nodules are formed from the middle cortex and they can be either determinate or indeterminate nodules depending on the growth condition (Ndoye et al., 1994; Fernandez-Lopez et al., 1998). In *Medicago*, inner cortical cells dedifferentiate, which results in the formation of nodule primordia that will subsequently differentiate into an indeterminate nodule (Xiao et al., 2014).

Release of rhizobia from ITs into nodule primordium cells occurs by formation of unwalled infection droplets. These are regions at the ITs in primordium/nodule cells that still contain the host membrane, but lack cell wall. Creation of such a cell-wall free interface allows rhizobia to be released into the cytoplasm and to become surrounded by a host membrane (Brewin et al., 1994; Bolanos et al., 2004; Brewin, 2004). The released rhizobium surrounded by the host membrane, is a transient organelle and is named symbiosome (Roth, 1989; Emerich and Krishnan, 2014). After bacteria have been released, they divide, differentiate and fill the host cells. The differentiated rhizobia are named bacteroids. In determinate nodules, rhizobia divide after time within the plant membrane and form symbiosomes with multiple bacteroids (Brewin, 2004). In these cases, bacteroids are morphologically similar to free-living bacteria. In indeterminate nodules, the symbiosomes contain a single bacteroid that markedly enlarges. The enlargement of bacteroids is correlated with endoreduplication of the bacteroids and a loss of the ability to return to the free-living bacterial state. This terminal differentiation is triggered by NCR peptides (Chapter 5) (Mergaert et al., 2003; Mergaert et al., 2006; Wang et al., 2010). Based on their morphology three distinct bacteroid morphotypes have been described. These are elongated bacteroids (E-morphotype), spherical bacteroids (S-morphotype) and unmodified bacteroids (U-morphotype) (Oono et al., 2010; Czernic et al., 2015).

The symbiotic interaction is set in motion by the exchange of signals between the two symbiotic partners. The host secretes among others flavonoids, which in many cases trigger the expression of rhizobial genes required for nodulation (*nod* genes). In the presence of flavonoids, NodD proteins activate the expression of the other *nod* genes (Djordjevic et al., 1987; Oldroyd and Downie, 2004; Jones et al., 2007). The proteins encoded by the *nod* genes are involved in the synthesis and secretion of Nod factors. Nod factors are lipochito-oligosaccharides (LCOs). Nod factors produced by different rhizobial species vary in their structure due to specific *nod* genes. For example, *nodC* determines the length of the chito-oligosaccharide backbone, and certain *nod* genes specify the type of substitutions at both ends of the molecule (Denarie et al., 1996). *Nod* genes are absent in some rhizobium species, for example *Bradyrhizobium* sp. ORS278 that induces nodules on *Aeschynomene* (Giraud et al., 2007; Bonaldi et al., 2010; Bonaldi et al., 2011). In these cases it is unclear which signal molecules trigger the nodulation process.

Nod factors are involved in induction of the early steps of nodulation. They are recognized by the receptors MtLYK3/MtNFP of *Medicago* (Limpens et al., 2003; Arrighi et al., 2006; Smit et al., 2007) or the orthologous LjNFR1/LjNFR5 of *Lotus* (Radutoiu et al., 2003; Radutoiu et al., 2007). They belong to the LysM domain-containing receptor-like kinase (LYKs) family, and the LysM motifs are involved specifically in perception of Nod factors. After perception of Nod factors, the symbiotic signaling network is activated that comprises a conserved set of genes encoding, a plasma membrane localized LRR-type receptor kinase (MtDMI2/LjSYMRK) (Ane et al., 2002; Stracke et al., 2002; Demchenko et al., 2004); several components in the nuclear envelope including a cation channel (MtDMI1/LjCASTOR/LjPOLLUX) (Ane et al., 2004; Charpentier et al., 2008) and subunits of the nuclear pore (LjNUP85/LjNUP133/LjNENA) (Kanamori et al., 2006; Saito et al., 2007; Groth et al., 2010), a nuclear localized complex of a calcium calmodulin-dependent protein kinase (CCaMK; MtDMI3/LjCCaMK) (Catoira et al., 2000; Levy et al., 2004; Gleason et al., 2006; Tirichine et al., 2007) and the transcription factor (MtIPD3/LjCYCLOPS) that is activated by CCaMK (Yano et al., 2008; Capoen et al., 2011; Ovchinnikova et al., 2011). The Nod factor receptor activation leads to calcium oscillations in the nucleoplasm. Calcium oscillations are decoded by CCaMK and activated CCaMK phosphorylates CYCLOPS that is a transcription factor for *NIN* (Schauser et al., 1999; Marsh et al., 2007; Tirichine et al., 2007; Yano et al., 2008; Heckmann et al., 2011). The induction of cortical cell division further requires the activation of the cytokinin-receptor (MtCRE1/LjLHK1) (Gonzalez-Rizzo et al., 2006; Tirichine et al., 2007). Further downstream in the signaling pathway, GRAS-type proteins (NSP1/NSP2) act as a transcriptional regulators contributing to the induction of genes required for nodule initiation (Kalo et al., 2005; Smit et al., 2005; Heckmann et al., 2006; Murakami et al., 2006).

3. Characteristics of determinate and indeterminate nodules

Mature nodules consist of a nitrogen-fixing central tissue and uninfected peripheral nodule tissues. These are the nodule endodermis, cortex, and parenchyma. In the latter the vascular bundles are located (Fig. 2) (Vandewiel et al., 1990; Brewin, 1991). The central tissue contains uninfected as well as enlarged infected plant cells.

The nodules formed by legumes are classified as either determinate or indeterminate (Hirsch, 1992; Sprent, 2001). A major difference between these two types is the life span of their meristem. Indeterminate nodules have a persistent meristem at their apex and obtain an elongated shape. In contrast, determinate nodules have a meristem at the periphery and this is only active at early stages of development, resulting in a spherical shape. The best characterized indeterminate nodules are those of for example *Medicago* and pea, and typical determinate nodules are formed by for example soybean, common bean, and *Lotus* (Hirsch, 1992).

Due to the difference in meristem persistence there are some additional characteristics that distinguish these two nodule types. The meristem of indeterminate nodules adds cells to the different nodule tissues in proximal direction. As a consequence, the central tissue shows a differentiation gradient with the youngest cells adjacent to the apical meristem and the oldest cells near the root attachment point. This zonation is indicated in Fig. 2. The meristem is called zone I. The subsequent zone is the infection zone (zone II). In the distal part of this zone, rhizobia are released, after which they divide and differentiate. The switch from infection to fixation zone (zone III) is a developmental switch associated with many changes after that. For example the *nif* genes are switched on, and amyloplasts accumulate at the periphery of the infected cells by which the start of the fixation zone can be easily recognized. The oldest zone is the senescence zone (zone IV) where bacteria are degenerated (Vasse et al., 1990). Further, mature tissues are absent at the nodule apex (Fig. 2). This is for example the case for the peripheral tissues, like endodermis and nodule parenchyma. In contrast, in determinate nodules these tissues completely surround the nodule central tissue (Fig. 2).

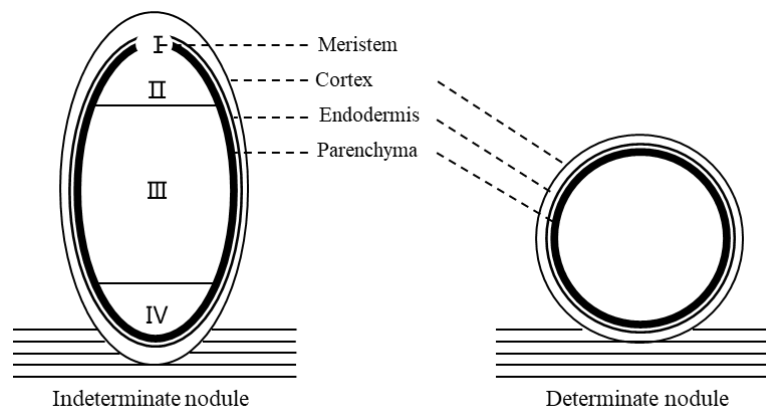


Fig. 2 Indeterminate and determinate nodules. The central tissue is surrounded by parenchyma, endodermis, and cortex. Indeterminate nodules originate from inner cortex and have a differential zonation: meristem (zone I), infection zone (zone II), nitrogen-fixing zone (zone III), and senescence zone (zone IV). Determinate nodules originate from the outermost or middle cortical cells and have no meristem at the apex.

In addition to the determinate and indeterminate nodule types, two other nodule types have been described. These are the lupinoid nodule formed by *Lupinus* species of the genistoids clade and the aeschynomene nodule formed by *Aeschynomene* species of the dalbergioids clade (Lotocka et al., 2000; Lavin et al., 2001). Lupinoid nodules are described as a subtype of indeterminate nodule (Lotocka et al., 2000). This nodule type is characterized by the presence of two lateral meristems, by which nodules encircle the roots (Golinowski et al., 1987). Another characteristic is that the infected cells remain mitotically active and the symbiosomes are equally distributed between the daughter cells (Gonzalez-Sama et al., 2004;

Fedorova et al., 2007). This mitotic activity might be facilitated by the absence of infection threads in these nodule cells (Lotočka et al., 2000; Gonzalez-Sama et al., 2004). Aeschynomenoid nodules are characterized by the fact that central infected tissue contains no uninfected interstitial cells. In some *Aeschynomene* species (e.g. *A. afraspera*), an out-growth at the top of nodules leads to indeterminate growth of the nodule (Bonaldi et al., 2011). This out-growth is derived from outer cortical cells containing large tubular structures with bacteria enclosed (Bonaldi et al., 2011). Lupinoid and aeschynomenoid nodules are proposed to be derived from an indeterminate ancestor and are considered to be a synapomorphy for the genistoids and dalbergioids groups, respectively (Lavin et al., 2001; Ardley et al., 2013).

4. Evolution of determinate and indeterminate nodules

The general accepted idea is that the determinate nodule type evolved from the indeterminate nodule type (Sprent, 2007). Knowledge concerning the distribution of the indeterminate and determinate nodule types within the Papilionoideae subfamily, when I started my research described in this thesis, is indicated in Fig. 1. The studied species in genistoids all form indeterminate nodules. Within the robinoids and dalbergioids, both determinate and indeterminate nodule types have been described. Several milletioids species form determinate nodules, an exception is *Tephrosia* forming indeterminate nodules. *Indigofera* from the indigoferoids is described as indeterminate nodule. All the studied species in the IRLC clade form indeterminate nodule. Together, this led to the hypothesis that the youngest ancestor of the Papilionoideae subfamily formed indeterminate nodules and the determinate nodule type evolved independently two times in the robinoids and milletioids, and once in dalbergioids clade (Doyle and Luckow, 2003; Sprent, 2007).

Nodules have been classified as indeterminate often based on the presence of dividing cells in mature nodules. In species of the IRLC (e.g. *Medicago*, pea), the meristem is formed at the primordium stage and is maintained throughout the life span of the nodule. This has not been studied in species of other clades. As some nodules that are classified as indeterminate nodules have characteristics of determinate nodules (Gehlot et al., 2012), I analyzed, in this thesis, the development of nodules of species belonging to the different clades of the Papilionoideae subfamily (Chapter 4 and Chapter 6).

5. Legume model systems

Although the Leguminosae family is the third-largest land plant family, only a few species are well studied at the molecular genetic level. This includes the model systems *Lotus* (*Lotus japonicus*) and *Medicago* (*Medicago truncatula*) (Barker et al., 1990; Handberg and Stougaard, 1992; Lohar et al., 2001; Pedrosa et al., 2002; Limpens et al., 2004; Sato et al., 2008; Young et al., 2011). These two species have all the characteristics required for in depth analyses using the tools of for example molecular biology, cellular biology and genetics. These two model plants have a relatively small genome, and their genomes have been

sequenced, providing insight into the evolution of rhizobial symbioses (Sato et al., 2008; Young et al., 2011). This is now extended with genome sequences of several accessions by which natural variation can be exploited. Further, good mutant collections have been created. Together, this has provided a strong basis to unravel molecular mechanism underlying for example Nod factor signalling, as well as nodule development and functioning. However, these well studied model legumes are temperate species, no species have been well studied that can survive under harsh abiotic conditions. In this thesis, *I. argentea*, which grows as pioneer vegetation in the Jizan desert (Saudi Arabia), has been selected to explore mechanisms used by legumes to survive under abiotic stress conditions.

References

- An , J.M., Kiss, G.B., Riely, B.K., Penmetsa, R.V., Oldroyd, G.E., Ajax, C., Levy, J., Debelle, F., Baek, J.M., Kalo, P., Rosenberg, C., Roe, B.A., Long, S.R., Denarie, J., and Cook, D.R. (2004). *Medicago truncatula DMI1* required for bacterial and fungal symbioses in legumes. *Science* 303(5662): 1364-1367.
- An , J.M., Levy, J., Thoquet, P., Kulikova, O., De Billy, F., Penmetsa, V., Kim, D.J., Debelle, F., Rosenberg, C., Cook, D.R., Bisseling, T., Huguet, T. and Denarie, J. (2002). Genetic and cytogenetic mapping of *DMI1*, *DMI2*, and *DMI3* genes of *Medicago truncatula* involved in Nod factor transduction, nodulation, and mycorrhization. *Molecular Plant-Microbe Interactions* 15(11): 1108-1118.
- Ardley, J.K., Reeve, W.G., O'Hara, G., Yates, R.J., Dilworth, M.J. and Howieson, J.G. (2013). Nodule morphology, symbiotic specificity and association with unusual rhizobia are distinguishing features of the genus *Listia* within the southern African crotalariaoid clade *Lotononis s.l.* *Annals of Botany* 112(1): 1-15.
- Arrighi, J.F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L.C., Mirabella, R., De Carvalho-Niebel, F., Journet, E.P., Gherardi, M., Huguet, T., Geurts, R., Denarie, J., Rouge, P., and Gough, C. (2006). The *Medicago truncatula* lysin [corrected] motif-receptor-like kinase gene family includes *NFP* and new nodule-expressed genes. *Plant Physiology* 142(1): 265-279.
- Barker, D.G., Bianchi, S., Blondon, F., Datt , Y., Duc, G., Essad, S., Flament, P., Gallusci, P., G n , G., Guy, P., Muel, X., Tourneur, J., D nari , J., and Huguet, T. (1990). *Medicago truncatula*, a model plant for studying the molecular genetics of the *Rhizobium*-legume symbiosis. *Plant Molecular Biology Reporter* 8(1): 40-49.
- Bolanos, L., Redondo-Nieto, M., Rivilla, R., Brewin, N.J., and Bonilla, I. (2004). Cell surface interactions of *Rhizobium* bacteroids and other bacterial strains with symbiosomal and peribacteroid membrane components from pea nodules. *Molecular Plant-Microbe Interactions* 17(2): 216-223.
- Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., Goormachtig, S., and Giraud, E. (2011). Nodulation of *Aeschynomene afraspera* and *A. indica* by photosynthetic *Bradyrhizobium sp.* strain ORS285: The Nod-dependent versus the Nod-independent symbiotic interaction. *Molecular Plant-Microbe Interactions* 24(11): 1359-1371.
- Bonaldi, K., Gherbi, H., Franche, C., Bastien, G., Fardoux, J., Barker, D., Giraud, E., and Cartieaux, F. (2010). The Nod factor-independent symbiotic signaling pathway: Development of *Agrobacterium rhizogenes*-mediated transformation for the legume *Aeschynomene indica*. *Molecular Plant-Microbe Interactions* 23(12): 1537-1544.
- Brewin, N.J. (1991). Development of the legume root nodule. *Annu Rev Cell Biol* 7: 191-226.
- Brewin, N.J. (2004). Plant cell wall remodelling in the *Rhizobium*-legume symbiosis. *Critical Reviews in Plant Sciences* 23(4): 293-316.
- Brewin, N.J., Rae, A.L., Perotto, S., Kannenberg, E.L., Rathbun, E.A., Lucas, M.M., Gunder, A., Bolanos, L., Kardailsky, I.V., and Wilson, K.E. (1994). Bacterial and plant glycoconjugates at the *Rhizobium*-legume interface. *Biochem Soc Symp* 60: 61-73.
- Callaham, D.A. and Torrey, J.G. (1981). The structural basis for infection of root hairs of *Trifolium repens* by *Rhizobium*. *Canadian Journal of Botany-Revue Canadienne De Botanique* 59(9): 1647-1664.
- Capoen, W., Oldroyd, G., Goormachtig, S., and Holsters, M. (2010). *Sesbania rostrata*: a case study of natural variation in legume nodulation. *New Phytologist* 186(2): 340-345.
- Capoen, W., Sun, J., Wysham, D., Otegui, M.S., Venkateshwaran, M., Hirsch, S., Miwa, H., Downie, J.A., Morris, R.J., Ane, J.M., and Oldroyd, G.E. (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.
- Catoira, R., Galera, C., De Billy, F., Penmetsa, R.V., Journet, E.P., Maillet, F., Rosenberg, C., Cook, D., Gough, C., and Denarie, J. (2000). Four genes of *Medicago truncatula* controlling components of a nod factor transduction pathway. *Plant Cell* 12(9): 1647-1666.
- Charpentier, M., Bredemeier, R., Wanner, G., Takeda, N., Schleiff, E., and Parniske, M. (2008). *Lotus japonicus* CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis. *Plant Cell* 20(12): 3467-3479.
- Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., Pierre, O., Fardoux, J., Chaintreuil, C., Nguyen, P., Gressent, F., Da Silva, C., Poulain, J., Wincker, P., Rofidal, V., Hem, S., Barriere, Q., Arrighi, J.F., Mergaert, P. and Giraud, E. (2015). Convergent evolution of endosymbiont differentiation in dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by Nodule-Specific Cysteine-Rich peptides. *Plant Physiology* 169(2): 1254-1265.

- Demchenko, K., Winzer, T., Stougaard, J., Parniske, M., and Pawlowski, K. (2004). Distinct roles of *Lotus japonicus* SYMRK and SYM15 in root colonization and arbuscule formation. *New Phytologist* 163(2): 381-392.
- Denarie, J., Debelle, F., and Prome, J.C. (1996). *Rhizobium* lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. *Annu Rev Biochem* 65: 503-535.
- Djordjevic, M.A., Redmond, J.W., Batley, M., and Rolfe, B.G. (1987). Clovers secrete specific phenolic compounds which either stimulate or repress nod gene expression in *Rhizobium trifolii*. *EMBO J* 6(5): 1173-1179.
- Doyle, J.J., and Luckow, M.A. (2003). The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiology* 131(3): 900-910.
- Emerich, D.W., and Krishnan, H.B. (2014). Symbiosomes: temporary moonlighting organelles. *Biochem J* 460(1): 1-11.
- Fedorova, E.E., De Felipe, M.R., Pueyo, J.J., and Lucas, M.M. (2007). Conformation of cytoskeletal elements during the division of infected *Lupinus albus* L. nodule cells. *J Exp Bot* 58(8): 2225-2236.
- Fernández-lópez, M., Goormachtig, S., Gao, M., D'Haese, W., Van Montagu, M., and Holsters, M. (1998). Ethylene-mediated phenotypic plasticity in root nodule development on *Sesbania rostrata*. *Proceedings of the National Academy of Sciences of the United States of America* 95(21): 12724-12728.
- Fields, S. (2004). Global nitrogen: Cycling out of control. *Environmental Health Perspectives* 112(10): A556-A563.
- Gage, D.J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol Mol Biol Rev* 68(2): 280-300.
- Gehlot, H.S., Panwar, D., Tak, N., Tak, A., Sankhla, I.S., Poonar, N., Parihar, R., Shekhawat, N.S., Kumar, M., Tiwari, R., Ardley, J., James, E.K., and Sprent, J.I. (2012). Nodulation of legumes from the Thar desert of India and molecular characterization of their rhizobia. *Plant and Soil* 357(1-2): 227-243.
- Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.C., Jaubert, M., Simon, D., Cartieaux, F., Prin, Y., Bena, G., Hannibal, L., Fardoux, J., Kojadinovic, M., Vuillet, L., Lajus, A., Cruveiller, S., Rouy, Z., Mangnot, S., Segurens, B., Dossat, C., Franck, W.L., Chang, W.S., Saunders, E., Bruce, D., Richardson, P., Normand, P., Dreyfus, B., Pignol, D., Stacey, G., Emerich, D., Vermeglio, A., Medigue C., and Sadowsky, M. (2007). Legumes symbioses: Absence of *Nod* genes in photosynthetic Bradyrhizobia. *Science* 316(5829): 1307-1312.
- Gleason, C., Chaudhuri, S., Yang, T., Munoz, A., Poovaiah, B.W., and Oldroyd, G.E. (2006). Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature* 441(7097): 1149-1152.
- Golinowski, W., Kopcinska, J., and Borucki, W. (1987). The morphogenesis of Lupine root nodules during infection by *Rhizobium lupini*. *Acta Societatis Botanicorum Poloniae* 56(4): 687-703.
- Gonzalez-Rizzo, S., Crespi, M., and Frugier, F. (2006). The *Medicago truncatula* CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with *Sinorhizobium meliloti*. *Plant Cell* 18(10): 2680-2693.
- Gonzalez-Sama, A., Lucas, M.M., De Felipe, M.R. and Pueyo, J.J. (2004). An unusual infection mechanism and nodule morphogenesis in white lupin (*Lupinus albus*). *New Phytologist* 163(2): 371-380.
- Groth, M., Takeda, N., Perry, J., Uchida, H., Draxl, S., Brachmann, A., Sato, S., Tabata, S., Kawaguchi, M., Wang, T.L., and Parniske, M. (2010). *NENA*, a *Lotus japonicus* homolog of *Sec13*, is required for Rhizodermal infection by arbuscular mycorrhiza fungi and Rhizobia but dispensable for cortical endosymbiotic development. *Plant Cell* 22(7): 2509-2526.
- Handberg, K., and Stougaard, J. (1992). *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. *Plant Journal* 2(4): 487-496.
- Heckmann, A.B., Lombardo, F., Miwa, H., Perry, J.A., Bunnewell, S., Parniske, M., Wang, T.L., and Downie, J.A. (2006). *Lotus japonicus* nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume. *Plant Physiology* 142(4): 1739-1750.
- Heckmann, A.B., Sandal, N., Bek, A.S., Madsen, L.H., Jurkiewicz, A., Nielsen, M.W., Tirichine, L., and Stougaard, J. (2011). Cytokinin induction of root nodule primordia in *Lotus japonicus* is regulated by a mechanism operating in the root cortex. *Molecular Plant-Microbe Interactions* 24(11): 1385-1395.
- Hirsch, A.M. (1992). Developmental biology of legume nodulation. *New Phytologist* 122(2): 211-237.
- Jones, K.M., Kobayashi, H., Davies, B.W., Taga, M.E., and Walker, G.C. (2007). How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. *Nature Reviews Microbiology* 5(8): 619-633.
- Kalò, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R.M., Hirsch, S., Jakab, J., Sims, S., Long, S.R., Rogers, J., Kiss, G.B., Downie, J.A., and Oldroyd, G.E. (2005). Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science* 308(5729): 1786-1789.

- Kanamori, N., Madsen, L.H., Radutoiu, S., Frantescu, M., Quistgaard, E.M., Miwa, H., Downie, J.A., James, E.K., Felle, H.H., Haaning, L.L., Jensen, T.H., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., and Stougaard, J. (2006). A nucleoporin is required for induction of Ca²⁺ spiking in legume nodule development and essential for rhizobial and fungal symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* 103(2): 359-364.
- Lavin, M., Herendeen, P.S., and Wojciechowski, M.F. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Syst Biol* 54(4): 575-594.
- Lavin, M., Pennington, R.T., Klitgaard, B.B., Sprent, J.I., De Lima, H.C., and Gasson, P.E. (2001). The dalbergioid legumes (Fabaceae): Delimitation of a pantropical monophyletic clade. *American Journal of Botany* 88(3): 503-533.
- Levy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.P., Ane, J.M., Lauber, E., Bisseling, T., Denarie, J., Rosenberg, C., and Debelle, F. (2004). A putative Ca²⁺ and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science* 303(5662): 1361-1364.
- Lewis, G.P. and Forest, F. (2005). Cercideae. Pp. 57-67 in: Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (eds.), *Legumes of the World*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Lewis, G.P., Schrire, B.D., Mackinder, B.A., Rico, L. and Clark, R. (2013). A 2013 linear sequence of legume genera set in a phylogenetic context: A tool for collections management and taxon sampling. *S. African J. Bot.* 89: 76-84. <https://doi.org/10.1016/j.sajb.2013.06.005>
- Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* 302(5645): 630-633.
- Limpens, E., Ramos, J., Franken, C., Raz, V., Compaan, B., Franssen, H., Bisseling, T., and Geurts, R. (2004). RNA interference in *Agrobacterium rhizogenes*-transformed roots of *Arabidopsis* and *Medicago truncatula*. *J Exp Bot* 55(399): 983-992.
- Lohar, D.P., Schuller, K., Buzas, D.M., Gresshoff, P.M., and Stiller, J. (2001). Transformation of *Lotus japonicus* using the herbicide resistance bar gene as a selectable marker. *J Exp Bot* 52(361): 1697-1702.
- Lotocka, B., Kopcinska, J., Gorecka, M., and Golinowski, W. (2000). Formation and abortion of root nodule primordia in *Lupinus luteus* L. *Acta Biologica Cracoviensia Series Botanica* 42(1): 87-102.
- LPWG, Legume Phylogeny Working Group (2017). A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. *TAXON* 66 (1): 44-77.
- 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. (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.
- Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., and Kondorosi, E. (2003). A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiology* 132(1): 161-173.
- Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., and Kondorosi, E. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* 103(13): 5230-5235.
- Murakami, Y., Miwa, H., Imaizumi-Anraku, H., Kouchi, H., Downie, J.A., Kawaguchi, M., and Kawasaki, S. (2006). Positional cloning identifies *Lotus japonicus* NSP2, a putative transcription factor of the GRAS family, required for NIN and ENOD40 gene expression in nodule initiation. *DNA Research* 13(6): 255-265.
- Ndoye, I., De Billy, F., Vasse, J., Dreyfus, B., and Truchet, G. (1994). Root nodulation of *Sesbania rostrata*. *J Bacteriol* 176(4): 1060-1068.
- Normand, P., and Bousquet, J. (1989). Phylogeny of nitrogenase sequences in *Frankia* and other nitrogen-fixing microorganisms. *Journal of Molecular Evolution* 29(5): 436-447.
- Novotny, V., Wang, X., Englande, A.J., Bedoya, D., Promakasikorn, L., and Tirado, R. (2010). Comparative assessment of pollution by the use of industrial agricultural fertilizers in four rapidly developing Asian countries. *Environment, Development and Sustainability* 12(4): 491-509.
- Oldroyd, G.E., and Downie, J.A. (2004). Calcium, kinases and nodulation signalling in legumes. *Nat Rev Mol Cell Biol* 5(7): 566-576.
- Oono, R., Schmitt, I., Sprent, J.I., and Denison, R.F. (2010). Multiple evolutionary origins of legume traits leading to extreme rhizobial differentiation. *New Phytologist* 187(2): 508-520.
- Ovchinnikova, E., Journet, E.P., Chabaud, M., Cosson, V., Ratet, P., Duc, G., Fedorova, E., Liu, W., Den Camp, R.O., Zhukov, V., Tikhonovich, I., Borisov, A., Bisseling, T., and Limpens, E. (2011). *IPD3* controls

- the formation of nitrogen-fixing symbiosomes in pea and *Medicago spp.* *Molecular Plant-Microbe Interactions* 24(11): 1333-1344.
- Pedrosa, A., Sandal, N., Stougaard, J., Schweizer, D., and Bachmair, A. (2002).** Chromosomal map of the model legume *Lotus japonicus*. *Genetics* 161(4): 1661-1672.
- Radutoiu, S., Madsen, L.H., Madsen, E.B., Felle, H.H., Umehara, Y., Gronlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., and Stougaard, J. (2003).** Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 425(6958): 585-592.
- Radutoiu, S., Madsen, L.H., Madsen, E.B., Jurkiewicz, A., Fukai, E., Quistgaard, E.M., Albrechtsen, A.S., James, E.K., Thirup, S., and Stougaard, J. (2007).** LysM domains mediate lipochitin-oligosaccharide recognition and Nfr genes extend the symbiotic host range. *EMBO J* 26(17): 3923-3935.
- Rae, A.L., Bonfanteasolo, P., and Brewin, N.J. (1992).** Structure and growth of infection threads in the legume symbiosis with *Rhizobium leguminosarum*. *Plant Journal* 2(3): 385-395.
- Roth LE, S.G. (1989).** Bacterium release into host cells of nitrogen-fixing soybean nodules: the symbiosome membrane comes from three sources. *European Journal of Cell Biology* 49(1): 13-23.
- Saito, K., Yoshikawa, M., Yano, K., Miwa, H., Uchida, H., Asamizu, E., Sato, S., Tabata, S., Imaizumi-Anraku, H., Umehara, Y., Kouchi, H., Murooka, Y., Szczyglowski, K., Downie, J.A., Parniske, M., Hayashi, M., and Kawaguchi, M. (2007).** NUCLEOPORIN85 is required for calcium spiking, fungal and bacterial symbioses, and seed production in *Lotus japonicus*. *Plant Cell* 19(2): 610-624.
- Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., Sasamoto, S., Watanabe, A., Ono, A., Kawashima, K., Fujishiro, T., Katoh, M., Kohara, M., Kishida, Y., Minami, C., Nakayama, S., Nakazaki, N., Shimizu, Y., Shinpo, S., Takahashi, C., Wada, T., Yamada, M., Ohmido, N., Hayashi, M., Fukui, K., Baba, T., Nakamichi, T., Mori, H., and Tabata, S. (2008).** Genome structure of the legume, *Lotus japonicus*. *DNA Research* 15(4): 227-239.
- Schauser, L., Roussis, A., Stiller, J., and Stougaard, J. (1999).** A plant regulator controlling development of symbiotic root nodules. *Nature* 402(6758): 191-195.
- Smit, P., Limpens, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C., and Bisseling, T. (2007).** Medicago LYK3, an entry receptor in rhizobial nodulation factor signaling. *Plant Physiology* 145(1): 183-191.
- Smit, P., Raedts, J., Portyanko, V., Debelle, F., Gough, C., Bisseling, T., and Geurts, R. (2005).** NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. *Science* 308(5729): 1789-1791.
- Sprent, J.I. (2001).** Formation, structure and function of nodules. *Nodulation in legumes*. J. I. Sprent. Kew, Royal Botanic Gardens: 35.
- Sprent, J.I. (2007).** Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytologist* 174(1): 11-25.
- Sprent, J.I., Ardley, J., and James, E.K., (2017).** Biogeography of nodulated legumes and their nitrogen fixing symbionts. *New Phytologist* 215: 40-56.
- Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Szczyglowski, K., and Parniske, M. (2002).** A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* 417(6892): 959-962.
- Tirichine, L., Sandal, N., Madsen, L.H., Radutoiu, S., Albrechtsen, A.S., Sato, S., Asamizu, E., Tabata, S., and Stougaard, J. (2007).** A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science* 315(5808): 104-107.
- Vandewiel, C., Scheres, B., Franssen, H., Vanlierop, M.J., Vanlammeren, A., Van Kammen, A., and Bisseling, T. (1990).** The early nodulin transcript *ENOD2* is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules. *EMBO J* 9(1): 1-7.
- Vasse, J., De Billy, F., Camut, S., and Truchet, G. (1990).** Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J Bacteriol* 172(8): 4295-4306.
- Wang, D., Griffiths, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., Bisseling, T., and Long, S.R. (2010).** A Nodule-Specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* 327(5969): 1126-1129.
- Wojciechowski, M.F., Lavin, M., and Sanderson, M.J. (2004).** A phylogeny of legumes (Leguminosae) based on analyses of the plastid matK gene resolves many well-supported subclades within the family. *American Journal of Botany* 91(11): 1846-1862.
- 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(18): 3517-3528.
- Yano, K., Yoshida, S., Muller, J., Singh, S., Banba, M., Vickers, K., Markmann, K., White, C., Schuller, B., Sato, S., Asamizu, E., Tabata, S., Murooka, Y., Perry, J., Wang, T.L., Kawaguchi, M., Imaizumi-Anraku,**

- H., Hayashi, M., and Parniske, M. (2008).** CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proceedings of the National Academy of Sciences of the United States of America* 105(51): 20540-20545.
- Young, N.D., Debelle, F., Oldroyd, G.E., Geurts, R., Cannon, S.B., Udvardi, M.K., Benedito, V.A., Mayer, K.F.X., Gouzy, J., Schoof, H., et al. (2011).** The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480(7378): 520-524.

Chapter 2

A technology platform for *Indigofera argentea*

Guiling Ren¹, Robin van Velzen¹, Arjan van Zeijl¹, Jacob Thomas², Carolien Franken¹, Tingting Xiao¹, Henk Franssen¹, Ahmed H. Alfarhan² and Ton Bisseling¹

¹ Laboratory of Molecular Biology, Plant Sciences Group, Wageningen University and Research Centre, Wageningen, The Netherlands

² Botany & Microbiology Dept., College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

Abstract

Legumes can establish a nitrogen fixing nodule symbiosis with rhizobia. *Indigofera argentea* establishes this nodule symbiosis under harsh desert conditions. Therefore this species is a good system to study the adaptations allowing nodulation under abiotic stress conditions. Here we did a series of observations and experiments to make *I. argentea* suitable for molecular studies. We show that *I. argentea* is diploid ($2n=16$) with a genome size of 690 Mbp, and a relative short seed-to-seed generation time of three months. The *de novo* transcriptome assembly presented here is the first large-scale molecular resource for *I. argentea*. Further, we developed a protocol for efficient *Agrobacterium rhizogenes*-mediated root transformation by which the basis for functional analysis of symbiosis related genes is created.

Introduction

Legumes belong to the Fabaceae. This is the third-largest land plant family containing more than 770 genera and 19,500 species (Lewis et al., 2005, 2013; LPWG, 2017). The majority of legume species are able to establish a very efficient nitrogen fixing nodule symbiosis with bacteria belonging to different genera, that are collectively named rhizobia (e.g. *Allorhizobium*, *Azorhizobium*, *Bradyrhizobium*, *Devosia*, *Ensifer*, *Mesorhizobium*, *Methylobacterium*, *Microvirga*, *Ochrobactrum*, *Phyllobacterium*, *Rhizobium*, *Neorhizobium*, *Pararhizobium*, *Shinella* (α -proteobacteria), and *Burkholderia*, *Cupriavidus* (β -proteobacteria) (Moulin et al., 2001; Gyaneshwar et al., 2011; Sprent et al., 2017; Andrews and Andrews, 2017). A few legume species are well studied at the molecular genetic level. These are the model systems *Lotus japonicus* (Lotus) and *Medicago truncatula* (Medicago). Further, some legume crops have been studied in some detail, examples are soybean, pea, and clover (Schmutz et al., 2010; Smykal et al., 2014; Zhukov et al., 2015; Alves-Carvalho et al., 2015; Webb et al., 2014). However, none of these legume species is adapted to severe environmental stress conditions, whereas many other legumes have this ability, including for example *Indigofera* spp., *Mimosa* spp. and *Tephrosia* spp., that are adapted to severe abiotic stress in desert areas (Hou et al., 2009; Gehlot et al., 2012).

To study mechanisms that control adaptation of for example nodule symbiosis to abiotic stress, we selected *Indigofera argentea*, which among others grows in desert areas of e.g. Jizan province of Saudi Arabia. Jizan is located in the south of Saudi Arabia at the border with Jemen. Jizan desert is subtropical and has an arid hot climate (Koppen-Geiger classification).

Indigofera is the third largest genus in the legume family, and it has about 750 species (Schrire et al., 2009). *I. argentea* is a perennial subshrub, about 50 cm high, with numerous branches. Its imparipinnate leaves are composed of 7-11 small leaflets. The flowers occur in a racemose inflorescence. The seedpods are 9-13 mm long and about 3 mm wide, in general they contain 5-7 seeds. *I. argentea* has a scattered population in Saudi Arabia, in the southwestern region it especially occurs in desert areas relatively close to the coast, for example at deserts near Baysh, Sabea and Abu Arish in Jizan Province. This region has a hot (up to 50°C) and humid climate and receives some rainfall (~100 mm/year) during winter time, whereas in summer it almost does not rain. *I. argentea* prefers to grow in well drained, sandy soil, along the edges of deep sand and shares the habitats of *Acacia ehrenbergiana*-*Capparis decidua* community or *Panicum turgidum*-*Dipterygium glaucum* community. In addition, we observed that it grows as a pioneer in former agricultural areas in Jizan desert and there we collected the plant material (Fig. 1).

We selected *I. argentea* to be able to study in the future the mechanisms by which this legume adapts to drought and heat. As a first step, we determined its genome size and chromosome

number. We showed that the genome size is rather small and it is diploid. Further, we developed an *A. rhizogenes*-mediated hairy root transformation protocol and assembled a *de novo* transcriptome, which was used to identify genes that are differentially regulated during nodule formation.

Results

Growth conditions

I. argentea was collected in Jizan desert at an area that had been used in the past for agriculture (Latitude 17.03.292; Longitude 042.38.139). In this area *I. argentea* is the most

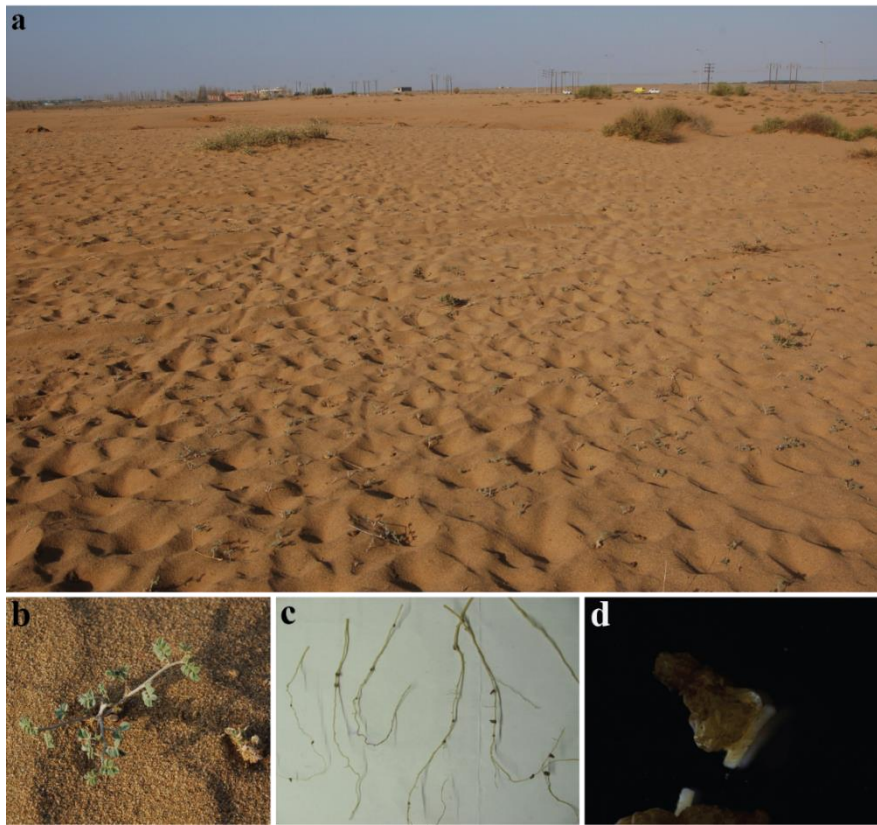


Fig. 1 *I. argentea* in Jizan desert. (a) *I. argentea* grows as a pioneer in former agricultural areas in Jizan desert; (b) *I. argentea* subshrub at vegetative stage; (c) Dry and brownish nodules formed on roots; (d) Nodules have an elongated shape.

abundant species (Fig. 1). The nutrient composition of the soil of this area was analysed (Table 1). Its pH is about 7. The content of Organic matter, total N (destructive), NH_4 , NO_2 , NO_3 , K and P is listed in Table 1 and shows that soil in Jizan desert is rather nutrient poor.

Table 1 Analyses of the soil sample, in red the detection limits of the analysis.

Method	A	B	C	D	E	F	G
Element	Nt	K	N-NH ₄	N-(NO ₃ +NO ₂)	C-elementary	C	P
Unit	[g/kg]	[mg/kg]	[mg/kg]	[mg/kg]	[mg/kg]	[g/kg]	[mg/kg]
detection limit	0.30	3.0	1.00	0.50	3.0	3.00	7.2
Sample: 13.4	0.0	33	1.1	0.7	11	2	2.1
							8

Method: (A) SFA-Nt/Pt destruction with $\text{H}_2\text{SO}_4\text{-H}_2\text{O}_2\text{-Se}$; (B) ICP-AES extraction in 0.01M CaCl_2 ; (C) SFA extraction with 0.01M CaCl_2 ; (D) SFA-Total Organic Carbon (TOC); (E) Spectrophotometer-Kurmes; (F) P-Olsen; (G) pH-meter.

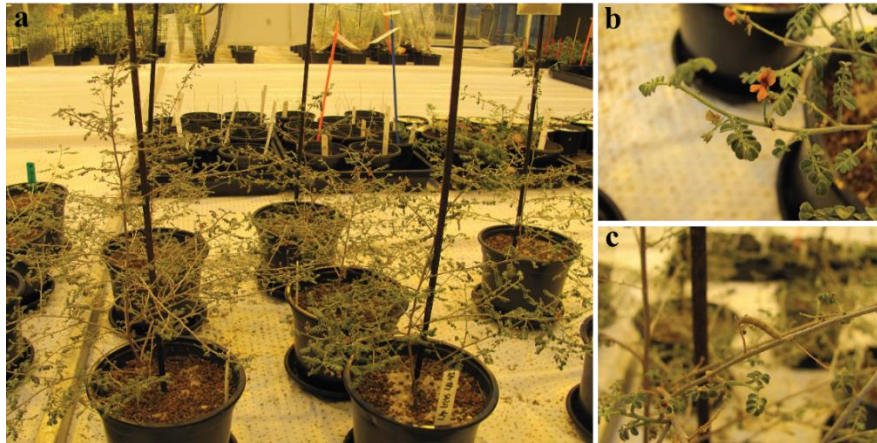


Fig. 2 *I. argentea* grown in the greenhouse. (a) *I. argentea* plants at reproduction stage; (b) Flowers are first formed about 2 months after sowing; (c) Pods are first formed about 3 months after sowing.

The growth conditions we established in the greenhouse are as follow: 28°C, >85% humidity, and 16 and 8 h of artificial light and darkness, respectively. It takes about 3 months from seed to seed under these growth conditions (see Materials and Methods; Fig. 2a). Flowering can be maintained after onset (Fig. 2b) at least for a year with continuous seed production and sufficient seeds can be harvested from a few plants. Matured yellow seeds have > 90% germination efficiency after scarification with concentrated sulfuric acid. Seeds kept at room temperature for three years still germinate with a high efficiency (>65%). Young seedlings are small enough to be grown 3 to 4 weeks on 9 cm Petri dishes containing Färhaeus medium (Fig. 4). Single seed derived lines were created from four plants that were generated from

seeds collected in Jizan. This resulted in 4 pure third generation lines. These are named Jizan-1, Jizan-2, Jizan-3, and Jizan-4.

Chromosome number and genome size

To determine the chromosome number of *I. argentea*, root tips were squashed and stained with acetic carmine (Belling, 1926; Fyad-Lameche et al., 2016). Cells at metaphase showed that *I. argentea* has 16 chromosomes (Fig. 3). Species of *Indigofera* are in general either diploid ($2n=2x=16$) such as *I. hochstetteri*, or tetraploid ($2n=4x=32$) such as *I. spicata*, or hexaploid ($2n=6x=48$) such as *I. heteranthera* (Frahm-Leliveld et al., 1962). Therefore *I. argentea* is most likely diploid ($2n=2x=16$). The 16 chromosomes are relatively large in size (total length 38.1 μm) and have a similar length.

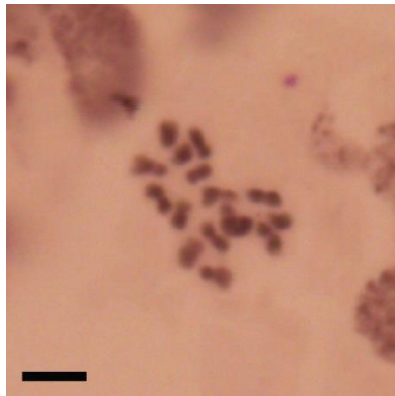


Fig. 3 Metaphase chromosomes of *I. argentea* after acetocarmine staining. Scale bar: 5 μm .

To determine the genome size, nuclei were isolated from young leaves. The genome size was determined by flow cytometry using *M. truncatula* (A17, 466 Mb), *Parasponia andersonii* (WU1, 563Mb), and soybean (1,115Mb) (Arumuganathan and Earle, 1991) as references. This showed that the genome size of *I. argentea* is approximately 690 Mb. The relative small genome size of *I. argentea* can facilitate molecular studies.

Hairy root transformation of *I. argentea* using *Agrobacterium rhizogenes*

To study the function of genes in root nodule formation, often composite plants with transgenic roots and non-transformed shoots are used. In general, such composite plants are created by *Agrobacterium rhizogenes*-mediated hairy root transformation. This method is well developed in model legumes (e.g. Medicago, Lotus), using the binary vector pRedRoot for selection of co-transformed roots (Limpens et al., 2004).

The susceptibility of *I. argentea* to *A. rhizogenes* was tested using three *A. rhizogenes* strains: RBL 1334, Arqua 1, and MSU440. The binary vector pRedRoot was introduced into these three strains. *I. argentea* seedlings were grown for 5-6 days at 28°C. At this stage, the first leaves are present (Fig. 4a). Seedlings were transformed by inoculating freshly cut hypocotyls with one of the three *A. rhizogenes* strains and the infection sites were kept in an environment with high humidity. *I. argentea* is a sub-tropical legume. Therefore, two growth temperatures 21°C and 28°C were tested for hairy root emergence (Fig. 4b, c). *I. argentea* grew better at 28°C and after two weeks on emergence medium, roots had developed (Fig. 4c-e). At 21°C only *A. rhizogenes* strain Arqua 1 induced a few hairy roots on 2 seedlings out of 30. Further, the seedlings were brownish and new leaves had not been formed (Fig. 4b). At 28°C, all three *A. rhizogenes* strains induced hairy roots on *I. argentea*, 23 seedlings out of 30 for strain Arqua 1 (76.7%), 27 out of 28 for strain RBL 1334 (96.4%), 6 out of 25 for strain MSU440 (24%). At least five homogeneously transformed roots (based on red fluorescence) were formed on each inoculated *I. argentea* seedling three weeks after transformation (Fig. 4d, e, arrow). The seedlings with transgenic roots induced by *A. rhizogenes* (Arqua 1 and RBL 1334) were inoculated with *Bradyrhizobium elkanii* (SA281) (Chapter 3) and grown on perlite at 28°C. One month after inoculation, nodules were formed on the transgenic roots (Fig. 4f, arrowhead). These results showed that *A. rhizogenes* (Arqua 1 and RBL 1334)-mediated hairy root transformation works well on *I. argentea*.

DR5::GUS auxin response pattern in *I. argentea* roots

To test the transformation procedure, we used DR5::GUS as a proof of principle. The auxin reporter DR5::GUS has been used to show auxin accumulation during root development in *Arabidopsis* (Sabatini et al., 1999), as well as root and nodule development in *Medicago* and soybean (Franssen et al., 2015; Turner et al., 2013). DR5::GUS was introduced into *I. argentea* by hairy root transformation using *A. rhizogenes* (Arqua 1). Longitudinal sections of transgenic roots showed that the DR5 promoter is expressed in the root meristem and columella cells and in lateral root primordia (Fig. 5a, b). The expression pattern and auxin responsiveness are consistent with what has been reported in *Arabidopsis*, soybean, and *Medicago*, indicating that the construct is suitable for monitoring auxin responses in *I. argentea* roots.

Identification of *I. argentea* genes with an enhanced expression in nodules and comparison with soybean and *Medicago*

To create a basis for comparative studies on *I. argentea* nodule expressed genes we conducted a *de novo* transcriptome analysis (see Materials and Methods). Then we selected a subset of the genes that are highly up-regulated in nodules compared to roots (> 10-fold) and a subset that are nodule-specific. Statistical analyses were performed using CLC Genomics. The

expression pattern in different tissues was assessed by statistical EDGE analysis based on common dispersion (P-value < 0.05, FDR corrected P-value < 0.05). 50 highly up-regulated

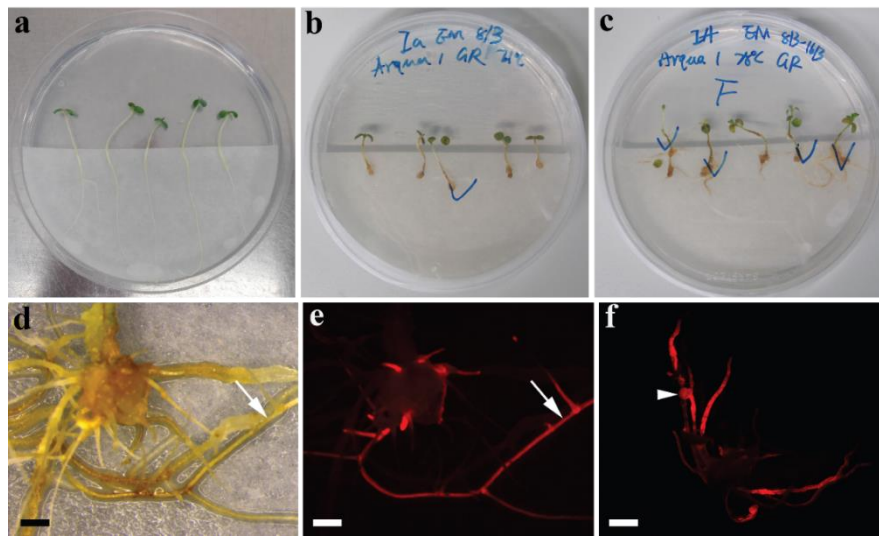


Fig. 4 *I. argentea* hairy root transformation. (a) 6-day-old seedlings grown at 28 °C, at this stage the first leaves are formed; (b) Seedlings inoculated with *A. rhizogenes* Arqua 1 growing at 21 °C for 7 days; (c) Seedlings inoculated with *A. rhizogenes* Arqua 1 growing at 28 °C for 7 days; (d, e) *I. argentea* root system with transgenic roots (arrow); (f) Nodule (arrowhead) on transgenic root three weeks after inoculation. Scale bars: 2mm.

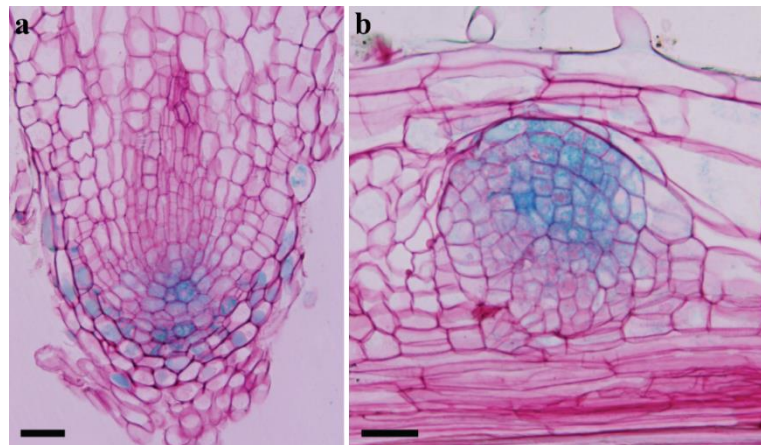


Fig. 5 DR5 promoter activity in *I. argentea* root. DR5::GUS transformed root show GUS activity especially in QC and part of the root cap (a), and lateral root primordia (b). Scale bars: 25 µm.

genes (> 10-fold) were selected and these genes were all highly expressed in nodules and roots (RPKM > 5 in nodule and root samples). Further, 145 nodule-specific genes were selected (RPKM = 0 in root, RPKM > 10 in nodule). These two subsets contain several well-known symbiosis genes, such as Leghemoglobin, glutamine synthetase, asparagine synthetase, *ENOD2* (Chapter 4), as well as NCRs (Chapter 5). Proteins encoded by these sets of genes were predicted based on their longest ORFs.

To obtain first indications that *I. argentea* nodules might have unique properties compared with for example the model legume Medicago and soybean, we first identified the closest homologs of these 195 genes in Medicago, soybean, *Phaseolus vulgaris*, *Cajanus cajan*, and Lotus genome protein databases (see Materials and Methods) by Geneious Custom BLAST at the protein level (default settings in Geneious R8). The best hits (identities >60%) from each species database were extracted for phylogenetic analyses, which were performed using MUSCLE Alignment and Geneious Tree Builder (Neighbor-Joining) in Geneious R8 with the default settings. Examples are shown in Fig. S1-7. Fig. S1 shows for example that Ia_c114718|m.98260 is part of a cluster in which all 5 other legumes have at least one homolog. The genes in such cluster might be orthologs, but will be named close homolog in this chapter. Fig. S2 and S3 show examples in which case Medicago or soybean close homolog, respectively, is not found. Whereas in Fig. S4, S5, and S6 examples are shown with multiple homologs in soybean or Medicago, respectively. In Fig. S7, Ia_c22267_g1_i1|m.7347 has no close homologs in Medicago and soybean. For the 195 *Indigofera* nodule genes we found 117 close homologs in at least Medicago or soybean. For these genes, it was studied whether they have a higher expression in nodules compared to roots in Medicago or soybean (Young et al., 2011; Roux et al., 2014; Schmutz et al., 2010).

13 close homologs of both soybean and Medicago had a remarkable different expression pattern as in both Medicago and soybean they were higher expressed in roots than in nodules (Table 2). 8 of these 13 genes are similar to genes that in other plant species have been reported to be upregulated under abiotic stresses. These are the isoflavone reductase-like protein in grapefruit induced in response to UV irradiation, in coffee during a stress-response in leaves, and in rice induced by gibberellic acid (Lers et al., 1998; Brandalise et al., 2009; Wen et al., 2010); vacuolar processing enzyme in radish involved in floral bud abortion under heat stress, and in *Malus hupehensis* and *Arabidopsis* in response to high temperature stress (Zhang et al., 2013; Su et al., 2015); cytochrome P450 induced by salt and mannitol treatments in apple (An et al., 2017). Further, the sulphate transporter gene, which can be induced by both sulphur starvation and mycorrhiza formation in Lotus (Giovannetti et al., 2014). These 13 nodule enhanced *I. argentea* genes (Table 2) might be an adaptation of the nodule to desert conditions and with the methods described in this chapter this can now be studied.

Table 2 The selected subset of genes those have a remarkable different expression pattern

Gene ID	Protein ID	Putative function	Fold change (Nodule/Root)	Medicago	Fold change (Nodule/Root)	soybean	Fold change (Nodule/Root)
Ia_c60158_g3_i2	Ia_c60158_g3_i2 m.72133	transmembrane ascorbate ferrireductase	Nodule specific	Medtr1g033390.1	0.14	Glyma07g27950	Root (low)
Ia_c35420_g1_i2	Ia_c35420_g1_i2 m.13628	MYND finger protein	Nodule specific	Medtr8g447290.1	0.26	Glyma08g25396/ Glyma08g25400/ Glyma15g29560	not found/ 0.33/ 0.13
Ia_c43412_g1_i4	Ia_c43412_g1_i4 m.20229	kinase interacting (KIP1-like) family protein	Nodule specific	Medtr7g108070.1	0.22	Glyma03g36740/ Glyma19g39380	0.08/ 0.94
Ia_c55655_g2_i4	Ia_c55655_g2_i4 m.47190	high affinity sulfate transporter type 1	Nodule specific	Medtr2g008470.1	0.18	Glyma08g19240/ Glyma15g05760	0.02/ Root
Ia_c114670_g1_i1	Ia_c114670_g1_i1 m.98187	cytochrome P450 family 78 protein	216.1	Medtr1g097220.1	0.07	Glyma20g31260	low
Ia_c61781_g5_i1	Ia_c61781_g5_i1 m.85729	cytochrome P450 family 71 protein	Nodule specific	Medtr4g104610.1/0.93/ Medtr4g089025.1/ not found/ Medtr4g089055.1/0.4/ Medtr4g089030.1 not found Medtr7g075310.1/0.21/	low/ 0.16/ Root/ Root	Glyma09g26430/ Glyma16g32010/ Glyma07g31390/ Glyma07g31380	
Ia_c60466_g3_i2	Ia_c60466_g3_i2 m.74368	vacuolar processing enzyme	Nodule specific	Medtr7g079140.1	not found	Glyma16g07190	low
Ia_c60466_g3_i1	Ia_c60466_g3_i1 m.74367	vacuolar processing enzyme	16.96	Medtr7g079140.1	not found	Glyma16g07190	low
Ia_c56248_g1_i5	Ia_c56248_g1_i5 m.49867	isoflavone reductase-like protein	Nodule specific	Medtr5g020760.1/0.85/ Medtr5g020800.1 0.38		Glyma01g37810/ Glyma11g07520/ Glyma11g07510/ 0.25/ Glyma1454s00200	0.09/ 0.06/ 0.25/ 0.77
Ia_c44377_g1_i1/ Ia_c44377_g1_i2	Ia_c44377_g1_i1 m.21396/ Ia_c44377_g1_i2 m.21397	oxidoreductase/ferric-chelate reductase	Nodule specific	Medtr1g100150.1/0.02/ Medtr1g100130.1 0.47		Glyma10g37600	Root
Ia_c116577_g1_i1	Ia_c116577_g1_i1 m.99025/ Ia_c116577_g1_i1 m.99026	lysine/ornithine decarboxylase	Nodule specific	Medtr3g114870.1/0.31/ Medtr3g114890.1 not found		Glyma06g02290/ Glyma04g02230	0.05/ 0.02
Ia_c47372_g1_i1	Ia_c47372_g1_i1 m.25705	transducin/WD40 repeat protein	Nodule specific	Medtr8g079820.1/0.78/ Medtr5g017800.1 1.03		Glyma17g20310/ Glyma05g10590	0.58/ 0.42
Ia_c60585_g2_i2	Ia_c60585_g2_i2 m.75216	ribonucleoside-diphosphate reductase, alpha chain	Nodule specific	Medtr3g009480.1/0.52/ Medtr8g092510.1 0.94		Glyma04g39500/ Glyma06g15350	0.52/ 0.32

Discussion

We describe a series of experiments and analysis to support that *I. argentea* has the potential to become a good model system to study nodule formation and functioning under abiotic stress conditions. We have shown that *I. argentea* is a diploid. Its genome size is rather small and about twice the size of *Medicago*. Another main criterion for the choice of this species is the short generation time, approximately three months, which makes it possible to finish several cycles within one year. Further, an efficient microsymbiont was isolated (Chapter 3). Efficient *Agrobacterium rhizogenes*-mediated hairy root transformation has been developed. This makes *I. argentea* now a system of which root/nodule properties can be studied at a molecular level.

In the *de novo* assembled transcriptome dataset, 13 genes have been identified that are highly up-regulated in nodules or are nodule-specifically expressed in *I. argentea*, whereas their closest homologs do not share this property in soybean and *Medicago*. The *I. argentea de novo* transcriptome assembled after Illumina sequencing is a good resource for studying *I. argentea* transcripts related to nodule symbiosis. Of special interest could be the mechanism by which secondary clusters of dividing cells are formed in nodules (Chapter 4) and mechanisms underlying nodule functioning under abiotic stress.

Materials and Methods

Plant material, seed germination and root inoculation

Seeds of *I. argentea* and sand were collected in Jizan desert, Saudi Arabia. Seeds were treated with 96% H₂SO₄ for 7 min and subsequently rinsed six times with distilled water, then seeds were sterilized with 4% commercial bleach for 10 min, rinsed seven times with sterilized distilled water followed by 3 h imbibition at room temperature in the dark. Sterilized seeds were transferred to 9 cm Petri dishes containing Färhaeus medium covered with filter paper at 4°C for 12 h in dark and 24 h in dark at 28°C. Subsequently, seeds were exposed to light and after 4-5 days, the seedlings were transferred to pots filled with river sand mixed with clay (Fig. 4a). After transfer of seedlings, nitrogen free Färhaeus medium was added. The seedlings were grown in the pots in the greenhouse for 3-4 days with no watering. Then the seedlings were inoculated with *B. elkanii* (SA281) (Chapter 3) and grown for another 2-3 days without watering. Subsequently, watering the plants twice every week with 100-200 ml for each pot and providing nitrogen free Färhaeus medium every two weeks.

Plasmids and vectors

Three *A. rhizogenes* strains, RBL 1334, Arqua 1, and MSU440 containing the pRedRoot vector (Limpens et al., 2004) were used for hairy root transformation of *I. argentea* under kanamycin selection (50 ug/ml). The plasmid DR5::GUS (Franssen et al., 2015) was

introduced into *A. rhizogenes* strain Arqua 1 by electro transformation. The strains were grown for 2 d at 28°C under spectinomycin selection (50 µg/ml).

Chromosome preparations and acetic carmine staining

Somatic metaphase chromosome preparations were made from actively growing root tips of *I. argentea* plants by squashing (Fyad-Lameche et al., 2016). Slides with chromosomes were frozen in liquid nitrogen, then cover slips were removed immediately with a razor blade. The chromosomes were stained with 1% solution of carmine in 45% acetic acid (Belling et al., 1926).

Genome size estimation

Young leaves were ground in liquid nitrogen, and the nuclei were collected as described (Van Velzen et al., 2017). Then the genome size was estimated by flow cytometry, using *M. truncatula* (A17, 466 Mb), *P. andersonii* (WU1, 563Mb), and soybean (1,115Mb) as references.

Tissue culture and *Agrobacterium rhizogenes*-mediated transformation

I. argentea hairy root transformation was done as described, with some modification (Limpens et al., 2004). Seeds were vernalized for 1 d at 4 °C and germinated at 28°C for 24 h in darkness (plates upside down). Seedlings were transferred to Petri dishes and grown at 28°C (16/8 h light/dark). After inoculated with one of the three *Agrobacterium* strains, the seedlings were co-cultivated for 5 d at 21°C (16/8 h light/dark) and subsequently transferred to Emergence medium. *Indigofera* plants were grown on Emergence medium at 21°C (16/8 h light/dark) inoculated with RBL 1334, Arqua 1 or MSU 44 for 2 weeks. The same was done at 28°C (16/8 h light/dark). Plants grown at 28°C were transferred to new Emergence medium with new filter paper after one week to avoid too extensive growth of *A. rhizogenes* strains. In this period, new roots were formed that are potentially co-transformed with the plasmid of the binary vector.

I. argentea was inoculated with *A. rhizogenes* strain Arqua 1 containing plasmid DR5::GUS following the same protocol.

Nodulation of *A. rhizogenes* transformed roots

Three weeks after transformation, composite *I. argentea* plants were transferred to sand-perlite (Maasmond-Westland, The Netherlands) saturated with Färhaeus medium (without Ca(NO₃)₂). Plants were starved for nitrate for 3 d at 28 °C (16/8 h light/dark). Then plants were inoculated with 500 µl of cultured *B. elkani* (SA281, OD₆₀₀=1.0) per plant and grown for 3-4 weeks at 28 °C (16/8 h light/dark).

Expression analysis and histochemical GUS staining

Plant tissues containing promoter-GUS fusions were stained as described (Franssen et al., 2015). Nodule and root sections of 5 µm thick were prepared with a RJ2035 microtome. Slides with sections were analysed with an AU5500B microscope equipped with a DFC425c camera (both Leica).

RNA isolation for sequencing

Uninfected roots, leaves, flowers, pods as well as nodules from *I. argentea* were frozen in liquid nitrogen and stored at -80°C until use. Total RNA, used for HiSeq 2000 sequence, was extracted as described (Van Velzen et al., 2017). After DNase I (Invitrogen) treatment, the integrity of total RNA was tested by gel electrophoresis. Isolated RNA samples were sent to BGI for sequencing.

De novo transcriptome assembly

Total RNA was isolated from different plant tissues including one-month old nodules, uninfected roots, leafs, flowers, and pods, and cDNA libraries were prepared. cDNA libraries were sequenced on an Illumina HiSeq 2000 platform, generating over 32 Gb paired-end data from about 400 million raw reads after removal of adapter and index sequences. These high quality reads were used to generate a primary assembly in Trinity software (k=31) (Haas et al., 2013; Grabherr et al., 2011). Trinity generated 162,573 linear contigs longer than 200 bases, grouped into 128,951 unique “transcripts”. The average length of the contigs was 996 bp, and the longest 15,875 bp (N50 of 1,905 bp) (Table S1).

Functional annotation

Functional annotation was performed at transcript and protein level using the Trinotate pipeline (<http://trinotate.github.io/>). Open reading frames (ORFs) were predicted using Transdecoder (Haas et al., 2013) and 100,206 protein sequences (longest protein 4,344 amino acids) were produced based on the longest ORFs. The annotation included homology search to NCBI nucleotide sequences (BLAST); the manually annotated and curated protein sequence database (SwissProt); protein domain identification by searching Protein Family database (PFAM) based on the profile hidden Markov models (HMMER); protein signal peptide prediction (signalP) and transmembrane domain prediction based on hidden Markov models (tmHMM); and leveraging various annotation databases (eggNOG/GO/Kegg databases). A total of 27,067 transcript sequences were assigned GO terms, including 21,849 with hits at the Biological Process level, 21,008 at the Cellular Component level and 23,245 at the Molecular Function level.

Phylogenetic analyses

Chapter 2

Protein datasets of soybean, Medicago, and *P. vulgaris* were downloaded from <https://phytozome.jgi.doe.gov/pz/portal.html>; *C. cajan* from <http://gigadb.org/dataset/100028>; and Lotus from <http://www.plantgdb.org/XGDB/phplib/download.php?GDB=Lj>. BLASTP and phylogenetic analysis were performed on Geneious R8 software (Kearse et al., 2012).

References

- Alves-Carvalho, S., Aubert, G., Carrere, S., Cruaud, C., Brochot, A.L., Jacquin, F., Klein, A., Martin, C., Boucherot, K., Kreplak, J., Da Silva, C., Moreau, S., Gamas, P., Wincker, P., Gouzy, J., and Burstin, J. (2015). Full-length *de novo* assembly of RNA-seq data in pea (*Pisum sativum* L.) provides a gene expression atlas and gives insights into root nodulation in this species. *Plant Journal* 84(1): 1-19.
- An, J.P., Li, R., Qu, F.J., You, C.X., Wang, X.F., and Hao, Y.J. (2017). Ectopic expression of an apple cytochrome P450 gene *MdCYP1* negatively regulates plant photomorphogenesis and stress response in *Arabidopsis*. *Biochemical and Biophysical Research Communications* 483(1): 1-9.
- Andrews, M., and Andrews, M.E. (2017). Specificity in Legume-Rhizobia Symbioses. *Int. J. Mol. Sci.* 18: 705.
- Arumuganathan, K., and Earle, E.D. (1991). Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* 9, 208–218.
- Belling, J. (1926). The iron-acetocarmine method of fixing and staining chromosomes. *Biological Bulletin* 50(2): 160-162.
- Brandalise, M., Severino, F., Maluf, M., and Maia, I. (2009). The promoter of a gene encoding an isoflavone reductase-like protein in coffee (*Coffea arabica*) drives a stress-responsive expression in leaves. *Plant Cell Reports* 28(11): 1699-1708.
- Frahm-Leliveld, J.A. (1962). Further observations on chromosomes in the genus *Indigofera* L. *Acta Botanica Neerlandica* 11: 201-208.
- Franssen, H.J., Xiao, T.T., Kulikova, O., Wan, X., Bisseling, T., Scheres, B., and Heidstra, R. (2015). Root developmental programs shape the *Medicago truncatula* nodule meristem. *Development* 142(17): 2941-2950.
- Fyad-Lameche, F.Z., Iantcheva, A., Siljak-Yakovlev, S., and Brown, S.C. (2016). Chromosome number, genome size, seed storage protein profile and competence for direct somatic embryo formation in Algerian annual *Medicago* species. *Plant Cell Tissue and Organ Culture* 124(3): 531-540.
- Gehlot, H.S., Panwar, D., Tak, N., Tak, A., Sankhla, I.S., Poonar, N., Parihar, R., Shekhawat, N.S., Kumar, M., Tiwari, R., Ardley, J., James, E.K., and Sprent, J.I. (2012). Nodulation of legumes from the Thar desert of India and molecular characterization of their rhizobia. *Plant and Soil* 357(1-2): 227-243.
- Giovannetti, M., Tolosano, M., Volpe, V., Kopriva, S., and Bonfante, P. (2014). Identification and functional characterization of a sulfate transporter induced by both sulfur starvation and mycorrhiza formation in *Lotus japonicus*. *New Phytologist* 204(3): 609-619.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q.D., Chen, Z.H., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology* 29(7): 644-652.
- Gyaneshwar, P., Hirsch, A.M., Moulin, L., Chen, W.M., Elliott, G.N., Bontemps, C., los Santos, P.E., Gross, E., dos Reis, F.B., Jr., Sprent, J.I., Young, J.P.W., and James, E.K. (2011). Legume-nodulating betaproteobacteria: diversity, host range and future prospects. *Molecular Plant-Microbe Interactions* 24: 1276-1288.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N., and Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8(8): 1494-1512.
- Hou, B.H., Wang, E.T., Li, Y., Jia, R.Z., Chen, W.F., Man, C.X., Sui, X.H., and Chen, W.X. (2009). Rhizobial resource associated with epidemic legumes in Tibet. *Microb Ecol* 57(1):69-81.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., and Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12): 1647-1649.
- Lers, A., Burd, S., Lomaniec, E., Droby, S., and Chalutz, E. (1998). The expression of a grapefruit gene encoding an isoflavone reductase-like protein is induced in response to UV irradiation. *Plant Molecular Biology* 36(6): 847-856.
- Lewis, G.P. and Forest, F. (2005). Cercideae. Pp. 57-67 in: Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (eds.), *Legumes of the World*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Lewis, G.P., Schrire, B.D., Mackinder, B.A., Rico, L. and Clark, R. (2013). A 2013 linear sequence of legume genera set in a phylogenetic context: A tool for collections management and taxon sampling. *S. African J. Bot.* 89: 76-84. <https://doi.org/10.1016/j.sajb.2013.06.005>

- 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*. *Journal of Experimental Botany* 55(399): 983-992.
- LPWG, Legume Phylogeny Working Group (2017). A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. *TAXON* 66 (1): 44-77.
- Moulin, L., Munive, A., Dreyfus, B., and Boivin-Masson, C. (2001). Nodulation of legumes by members of the beta-subclass of Proteobacteria. *Nature* 411(6840): 948-950.
- Robinson, N.J., Procter, C.M., Connolly, E.L., and Guerinot, M.L. (1999). A ferric-chelate reductase for iron uptake from soils. *Nature* 397(6721): 694.
- Roux, B., Rodde, N., Jardinaud, M., Timmers, T., Sauviac, L., Cottret, L., Carrère, S., Sallet, E., Courcelle, E., Moreau, S., Debellé, F., Capela, D., Carvalho-Niebel, F., Gouzy, J., Bruand, C., and Gamas, P. (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant Journal* 77(6): 817-837.
- Sabatini, S., Beis, D., Wolkenfelt, H., Murfett, J., Guilfoyle, T., Malamy, J., Benfey, P., Leyser, O., Bechtold, N., Weisbeek, P., and Scheres, B. (1999). An auxin-dependent distal organizer of pattern and polarity in the *Arabidopsis* root. *Cell* 99(5): 463-472.
- Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J.X., Mitros, T., Nelson, W., Hyten, D.L., Song, Q.J., Thelen, J.J., Cheng, J.L., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463(7278): 178-183.
- Schrire, B.D., Lavin, M., Barker, N.P., and Forest, F. (2009). Phylogeny of the tribe Indigofereae (Leguminosae-Papilionoideae): Geographically structured more in succulent-rich and temperate settings than in grass-rich environments. *Am J Bot* 96(4): 816-852.
- Smykal, P. (2014). Pea (*Pisum sativum* L.) in biology prior and after Mendel's discovery. *Czech Journal of Genetics and Plant Breeding* 50(2): 52-64.
- Sprent, J.I. (2007). Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytol* 174(1): 11-25.
- Sprent, J.I., Ardley, J., and James, E.K., (2017). Biogeography of nodulated legumes and their nitrogen fixing symbionts. *New Phytologist* 215: 40-56.
- Su, Q., Ran, K., Men, X.J., Zhang, W.W., Fan, S.L., Yan, L.J., and Yang, H.Q. (2015). Response of vacuolar processing enzyme in *Malus hupehensis* and *MhVPEy*-overexpressing *Arabidopsis* to high temperature stress. *Acta Physiologiae Plantarum* 37(4): 1-11.
- Turner, M., Nizampatnam, N.R., Baron, M., Coppin, S., Damodaran, S., Adhikari, S., Arunachalam, S.P., Yu, O., and Subramanian, S. (2013). Ectopic expression of miR160 results in auxin hypersensitivity, cytokinin hypersensitivity, and inhibition of symbiotic nodule development in soybean. *Plant Physiology* 162(4): 2042-2055.
- Van Velzen R., Holmer, R., Bu, F., Rutten, L., Van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D., Roswanjaya, Y.P., Wardhani, T.A.K., et al. (2017). Parallel loss of symbiosis genes in relatives of nitrogen-fixing non-legume *Parasponia*. *bioRxiv* (online Jul. 28, 2017).
- Webb, K.J., Cookson, A., Allison, G., Sullivan, M.L., and Winters, A.L. (2014). Polyphenol oxidase affects normal nodule development in red clover (*Trifolium pratense* L.). *Frontiers in Plant Science* 5: 1-14.
- Wen, F.P., Zhang, Z.H., Bai, T., Xu, Q., and Pan, Y.H. (2010). Proteomics reveals the effects of gibberellic acid (GA 3) on salt-stressed rice (*Oryza sativa* L.) shoots. *Plant Science* 178(2): 170-175.
- Young, N.D., Debelle, F., Oldroyd, G.E., Geurts, R., Cannon, S.B., Udvardi, M.K., Benedito, V.A., Mayer, K.F.X., et al. (2011). The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature* 480(7378): 520-524.
- Zhang, J., Li, Q.F., Huang, W.W., Xu, X.Y., Zhang, X.L., Hui, M.X., Zhang, M.K., and Zhang, L.G. (2013). A vacuolar processing enzyme *RsVPEI* gene of radish is involved in floral bud abortion under heat stress. *International journal of molecular sciences* 14(7): 13346-13359.
- Zhukov, V.A., Zhernakov, A.I., Kulaeva, O.A., Ershov, N.I., Borisov, A.Y., and Tikhonovich, I.A. (2015). *De novo* assembly of the pea (*Pisum sativum* L.) nodule transcriptome. *International Journal of Genomics*. Article ID 695947, 11 pages.

Supplementary Figures

Table S1 Transcriptome assembly metrics

Counts of transcripts, etc.	
Total trinity 'genes':	128,951
Total trinity transcripts:	162,573
GC content (%):	39.09
Statistics based on ALL transcript contigs:	
Contig N50:	1,905
Median contig length:	502
Average contig length:	996.47
Total assembled bases:	161,998,401
Statistics based on ONLY LONGEST ISOFORM per 'GENE':	
Contig N50:	1,434
Median contig length:	406
Average contig length:	786.6
Total assembled bases:	101,432,503

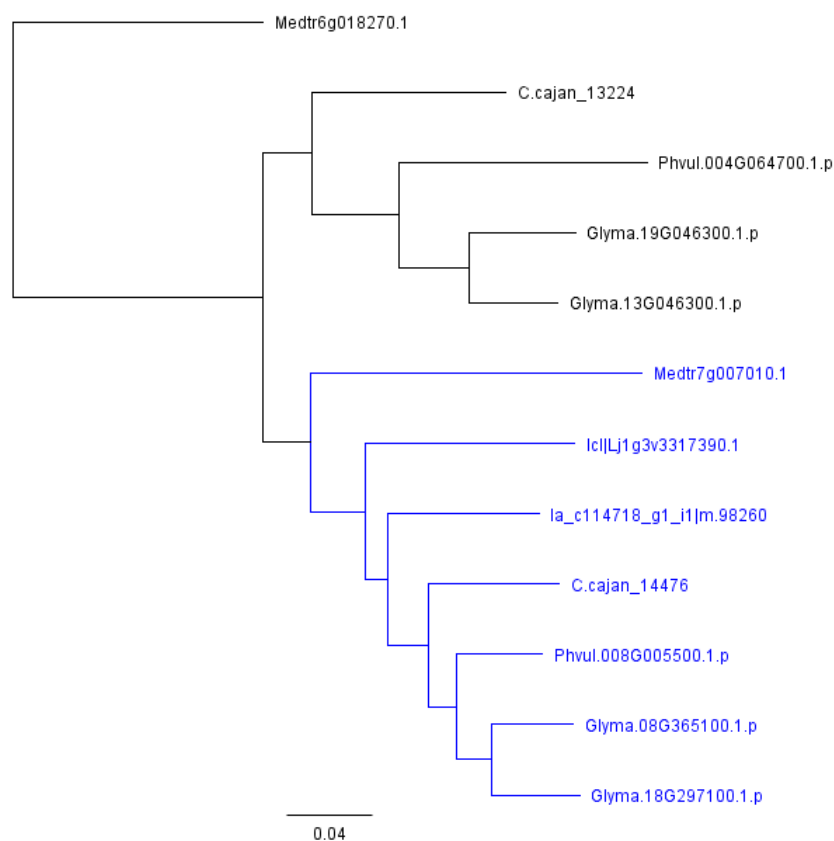


Fig. S1 Phylogenetic trees which shows the close homologs of Ia_c114718|m.98260 in Medicago (Medtr7g007010.1.p) and soybean (Glyma.08G365100.1.p, Glyma.18G297100.1.p) (in blue).

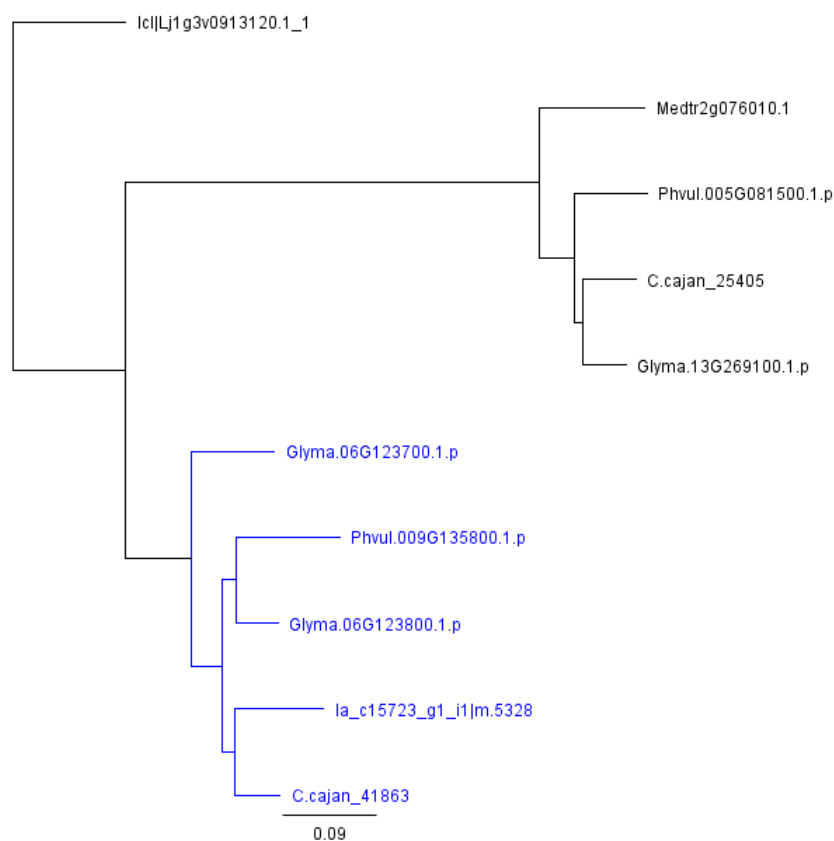


Fig. S2 Phylogenetic trees which showed the close homolog of Ia_c15723_g1_i1|m.5328 in soybean (Glyma.06G123800.1.p, Glyma.06G123700.1.p) (in blue), while no close homolog present in Medicago.

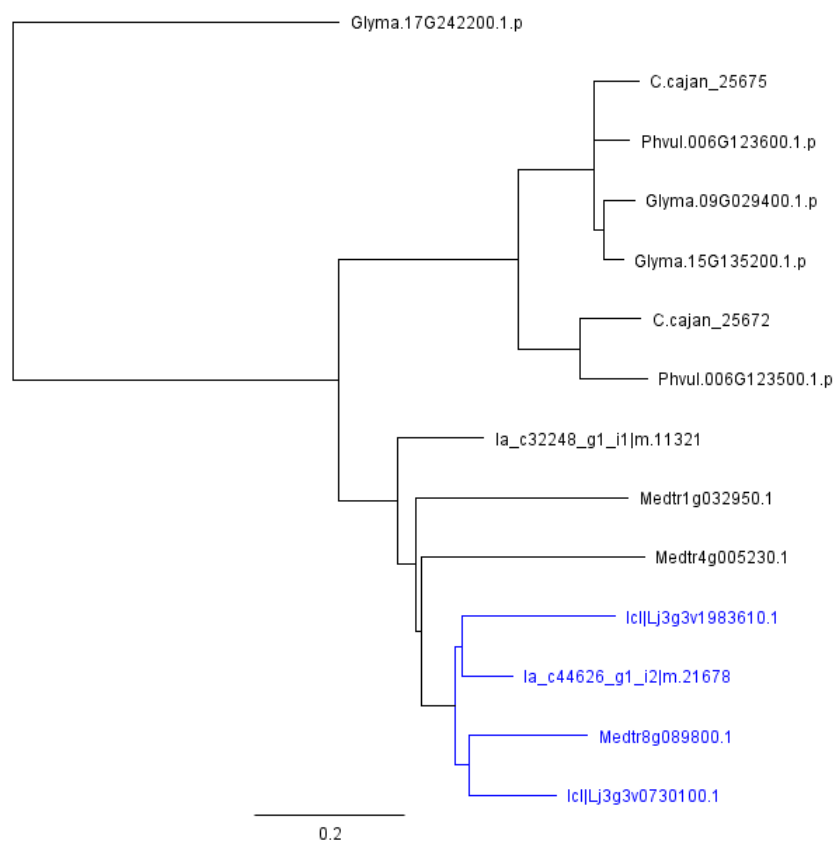


Fig. S3 Phylogenetic trees which shows the close homolog of la_c44626_g1_i2|m.21678 (in blue) in Medicago (Medtr8g089800.1), while no close homolog is present in soybean.

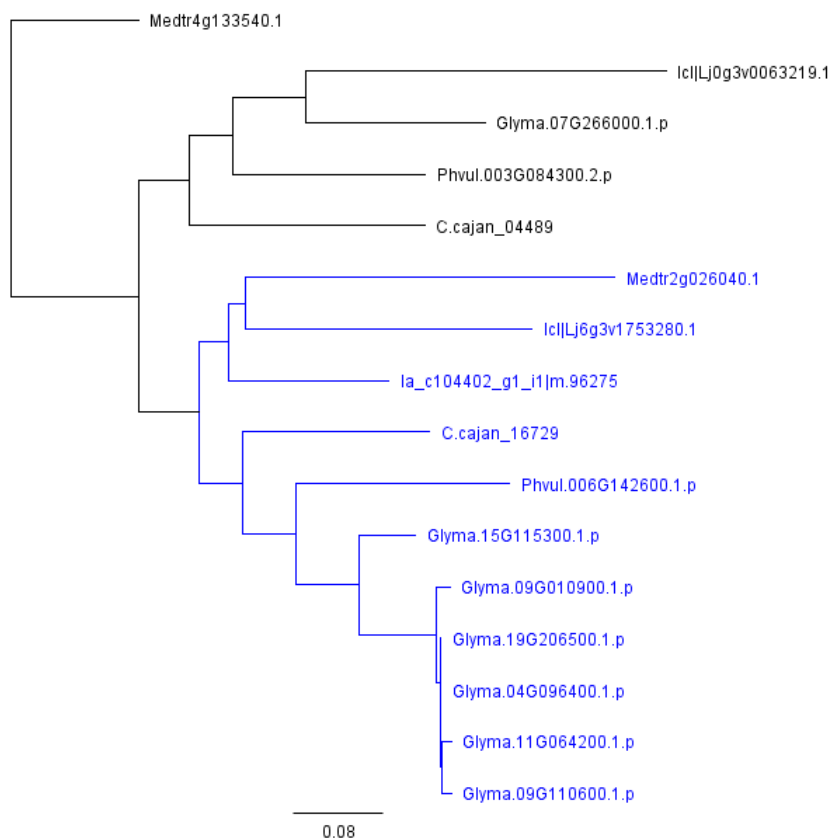


Fig. S4 Phylogenetic trees which shows the close homolog of *Ia_c104402_g1_i1|m.96275* in *Medicago* (*Medtr2g026040.1*) and multiple close homologs in soybean (*Glyma.15G115300.1.p*, *Glyma.09G010900.1.p*, *Glyma.11G064200.1.p*, *Glyma.09G110600.1.p*, *Glyma.19G206500.1.p*, *Glyma.04G096400.1.p*) (in blue)..

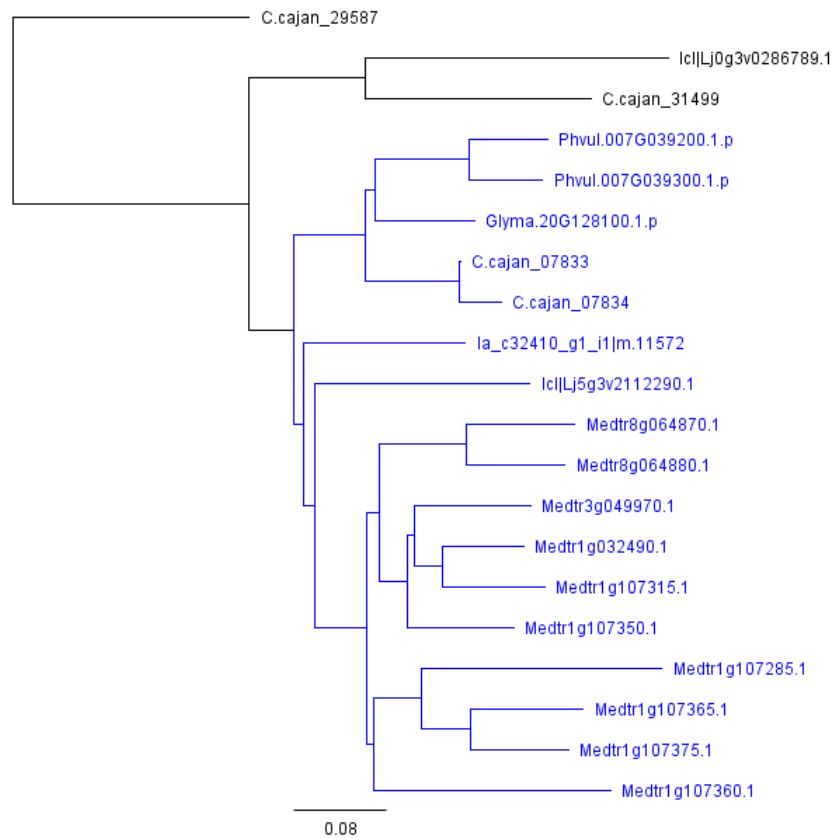


Fig. S5 Phylogenetic trees which shows the close homolog of Ia_c32410_g1_i1|m.11572 in soybean (Glyma.20G128100.1.p) and multiple close homologs in Medicago (Medtr8g064870.1, Medtr8g064880.1, Medtr3g049970.1, Medtr1g032490.1, Medtr1g107315.1, Medtr1g107350.1, Medtr1g107360.1, Medtr1g107285.1, Medtr1g107365.1, Medtr1g107375.1) (in blue).

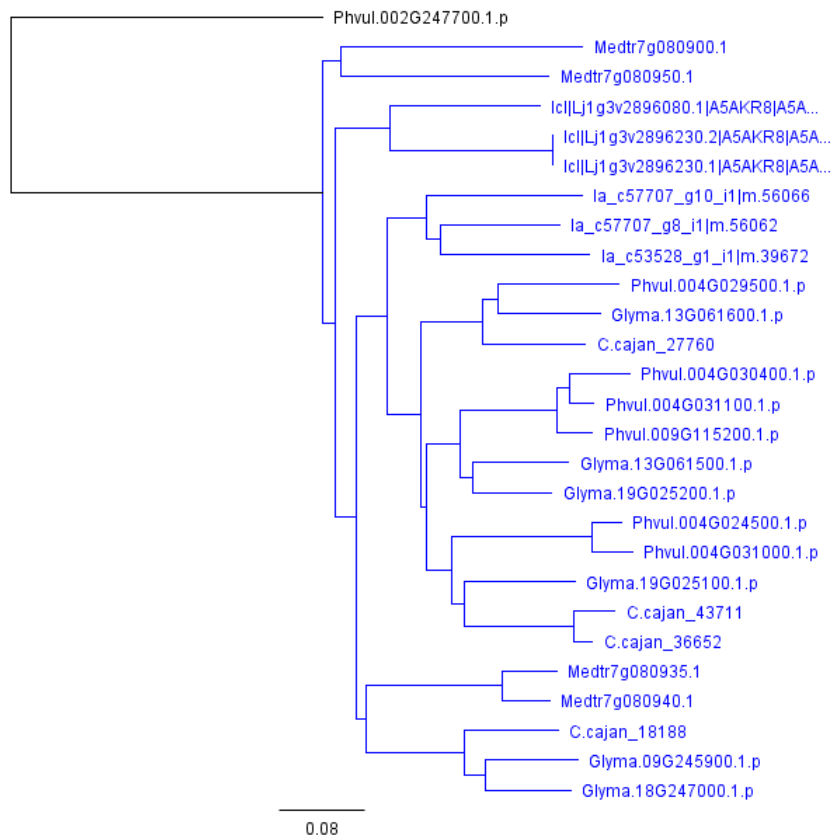


Fig. S6 Phylogenetic trees which shows the multiple close homologs of la_c57707_g10_i1|m.56066 in soybean and Medicago (in blue).

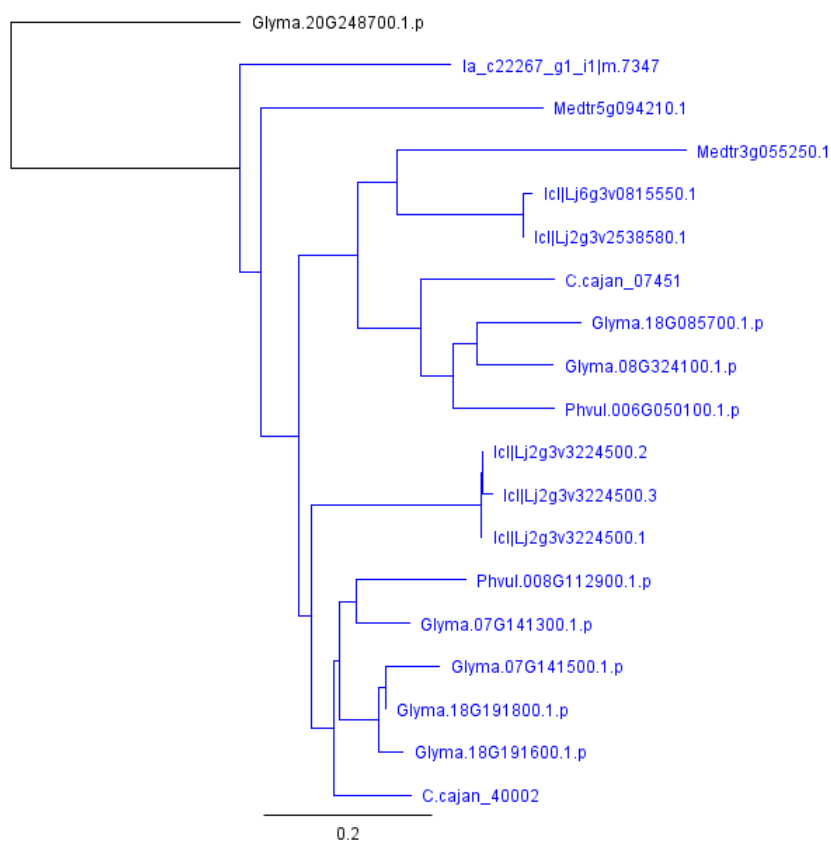


Fig. S7 Phylogenetic trees which shows that Ia_c22267_g1_i1|m.7347 has no close homologs in Medicago and soybean (in blue).

Chapter 3

Variation in infectiveness of rhizobium strains nodulating the promiscuous legume *Indigofera argentea*

Guiling Ren, Gijs Selten, Robin van Velzen, Xu Cheng, Arjan van Zeijl, Carolien Franken,
Henk Franssen, Ton Bisseling, and Rene Geurts

Laboratory of Molecular Biology, Plant Sciences Group, Wageningen University and
Research Centre, Wageningen, The Netherlands

Abstract

Legumes can establish a nitrogen fixing symbiosis with a diverse range of α -, β -, and γ -proteobacteria that are collectively named rhizobium. The host range of a particular legume species can vary from broad to very narrow. We investigated the diversity and infectiveness of rhizobia associated with the legume *Indigofera argentea*, a pioneer species in the coastal Jizan desert of Saudi Arabia. This showed that *I. argentea* is promiscuous and can form nodules with a diverse range of rhizobia, including different *Bradyrhizobium*, *Ensifer* and *Microvirga* species. For the latter this is the first report that a species of this genus can nodulate *Indigofera* spp.. Competition studies revealed that a single *Bradyrhizobium* species, represented by strain SA340, is most infective, as it occupies >50% of all nodules formed in mixed inoculum experiments. Nodules formed by *Bradyrhizobium* sp. SA340 display a cytoarchitecture that is in line with a compatible interaction. Further, this strain has profound growth promoting effects. Taken-together, we conclude that despite the promiscuous nature of *I. argentea*, it is predominantly nodulated by strains with a high infectiveness.

Introduction

The Fabaceae (legume family) encompasses over 19,500 species, of which the vast majority is able to establish a nitrogen fixing nodule symbiosis with a taxonomically diverse group of α - and β -proteobacterial species. These microbial symbionts are collectively known as rhizobium and include species of at least 16 genera; namely *Allorhizobium*, *Azorhizobium*, *Bradyrhizobium*, *Devosia*, *Ensifer* (*Sinorhizobium*), *Mesorhizobium*, *Methylobacterium*, *Microvirga*, *Ochrobactrum*, *Phyllobacterium*, *Rhizobium*, *Neorhizobium*, *Pararhizobium*, *Shinella* (α -proteobacteria), and *Burkholderia* and *Cupriavidus* (β -proteobacteria) (Moulin et al., 2001; Gyaneshwar et al., 2011; Sprent et al., 2017; Andrews and Andrews, 2017). Although there have been a few reports of rhizobial γ -proteobacteria (e.g. *Pseudomonas*), these have not been confirmed (Gyaneshwar et al., 2011; Moulin et al., 2015). The molecular nature of the specificity of this symbiosis has been studied mainly in two legume models, *Medicago truncatula* and *Lotus japonicus*, and a few crop species; e.g. soybean (*Glycine max*) and pea (*Pisum sativum*). However, all these cases represent species with a narrow microbial host range. In contrast, many legume species have a much broader host range, and can be nodulated with a diverse range of rhizobial species belonging to different taxonomic families. Here we studied *Indigofera argentea* to obtain insight in its host range, and the infectivity of compatible rhizobial species.

Indigofera is the third largest legume genus, comprising ~750 species divided over 4 main subclasses named according to their distribution centres; Palaeotropical (occurring in tropical Africa and Asia), Pantropical, Cape, and Tethyan clades (Schrire, 2005; Schrire et al., 2009). *Indigofera* is the largest genus within the tribe Indigofereae (Schrire et al., 2009). Many *Indigofera* species are adapted to drought and desert conditions, among which is *I. argentea*. *I. argentea* represents an early-divergent succulent biome species in the Palaeotropical subclade, which can be found in the Saharo-Sindian region (Schrire et al., 2009). It is a perennial subshrub that grows as pioneer vegetation in scattered populations in well drained, sandy soils, either alone, or together with other pioneer plant species like *Acacia ehrenbergiana*, *Capparis decidua*, *Panicum turgidum* and/or *Dipterygium glaucum* community. Also, *I. argentea* has a certain resilience to salt stress, and can grow as weed on salinized agricultural fields.

Studies in *M. truncatula*, *L. japonicus* and pea have revealed that their rhizobial host range is, at least in part, constrained by specificity in symbiotic signalling. Rhizobial-synthesized lipo-chitoligosaccharides (LCOs, also named Nodulation (Nod) factors) are perceived by a plant encoded heterodimer complex of LysM domain containing receptor like kinases (LysM-RKs). Upon perception a signalling cascade is induced that is essential for nodule organogenesis and bacterial infection. Nod factors carry various species-specific substitutions, and this structural variation is thought to play an important role in the specificity of the interaction (Lerouge et al., 1990; Perret et al., 2000; Radutoiu et al., 2007).

For example, by comparing *L. japonicus* to *L. filicaulis* it was shown that a differently positioned acetyl group on the Nod factor glucosamine backbone is discriminated by a single amino acid difference in a LysM domain of the NFR5 Nod factor receptor (Radutoiu et al., 2007). Similarly, it was found in pea, that ecotypes differ in host range that can be correlated to specificity in Nod factor perception (Geurts et al., 1997). Taken-together, this suggests that coevolution of plant LysM-RK receptors and rhizobium Nod factor structure may be an important driver in host specificity.

In addition to Nod factor specificity, other factors can affect the legume host range; e.g. perception of rhizobial surface polysaccharides, such as lipopolysaccharides, exopolysaccharides and cyclic glucans, and/or effector proteins secreted by the microsymbiont. Genetic studies in legumes start to uncover the mode of action of these inhibitory mechanisms. For example, in *L. japonicus*, a LysM-type RK has been found, named EPR3, that restricts entry of a *Mesorhizobium loti* strain with mutated exopolysaccharides (Kawaharada et al., 2015; Kawaharada et al., 2017), whereas in soybean cloning of *Rj* genes revealed that pathogenesis-related responses can cause host-range restriction (Hayashi et al., 2014; Yang et al., 2010).

As most molecular genetic studies have been performed on legume species with a restricted rhizobial host range, there is a lack of knowledge on the mechanisms that underlie symbiotic promiscuity. A wide range of legumes are defined as promiscuous, including *Indigofera* spp.. (Perret et al., 2000; Lira et al., 2015). For example, the microsymbionts of *Indigofera tinctoria* includes rhizobial species of the genera *Rhizobium*, *Ensifer*, *Bradyrhizobium*, *Cupriavidus* and *Pseudoalteromonas*, which represent three classes of the proteobacteria phylum (Leelahawong et al., 2010). We investigated the diversity and infectiveness of rhizobia associated with *I. argentea*. This confirmed that *I. argentea* is highly promiscuous and can form nodules with a diverse range of rhizobia, including different *Bradyrhizobium*, *Ensifer* and *Microvirga* species. Competition studies however, demonstrated strain specificity. Taken-together, we conclude that despite the promiscuous nature of *I. argentea*, it is predominantly nodulated by *Bradyrhizobium* sp. strain SA340 that has a high symbiotic infectiveness.

Results

Isolation and characterization of Jizan desert rhizobia that nodulate *I. argentea*

To get insight in the diversity of rhizobia that nodulate *I. argentea*, plants were grown in the greenhouse and “inoculated” with Jizan desert sand (see Material and Methods). After four weeks, roots were covered with pink nodules (Fig. 1a-c). Sectioning revealed a nodule cytoarchitecture similar as reported for legume nodules, including a large central tissue consisting of cells filled with symbiosomes (Fig. 1).

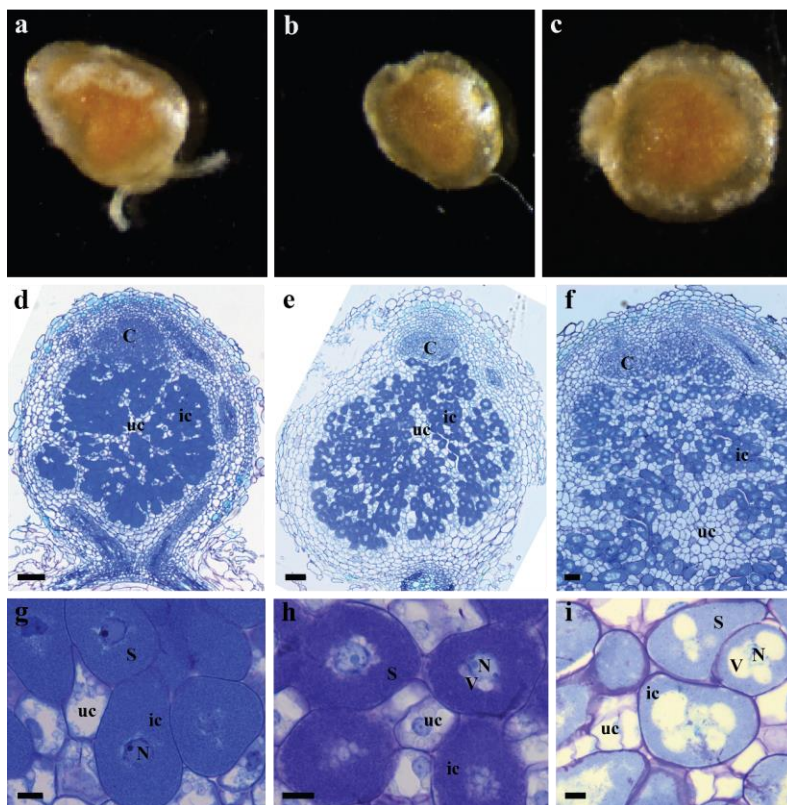


Fig. 1: Nodule morphology and longitudinal section of nodules induced by *B. genosp* SA340 (a, d, g), *E. kostiense* SA331 (b, e, h), and *Microvirga* sp. SA342 (c, f, i). (g-i) Zoom in of nodule sections in d-f to show the vacuole in the infected cells. C: cluster of secondary dividing infected cells; S: symbiosome; N: nucleus; V: vacuole; ic: infected cells; uc: uninfected cells; Bars: a-c, 150 μ m; d-f, 15 μ m.

From 55 surface sterilized nodules rhizobia were isolated. Morphologically the isolated strains had a different appearance, varying from globose to rod-shaped. Sequencing the 16S rDNA locus and subsequent BLAST analysis revealed that the strains represent three distinct genera: *Bradyrhizobium* (38 strains), *Ensifer* (9 strains), and *Microvirga* (8 strains). As *Microvirga* spp. was not reported as a microsymbiont of *Indigofera* species, it was determined whether the isolated strains contain the *nifH* gene encoding a subunit of the nitrogenase enzyme complex. This showed to be the case (Fig. 2). BLAST analysis of the *Microvirga nifH* sequence showed that the gene is most similar to the *nifH* gene from

Mesorhizobium temperatum (Fig. 2), suggesting an event of horizontal gene transfer between *Mesorhizobium* and *Microvirga*.

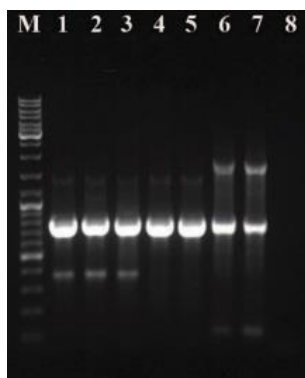


Fig. 2 Amplification of *nifH* gene from *Bradyrhizobium*, *Ensifer*, or *Microvirga* strains. (1-3), *Bradyrhizobium* strains SA328, SA314, SA333, (4-5), *Ensifer* strains SA331, SA222, (6-7), *Microvirga* strains SA342, SA348, (8), MQ negative control. SA222 was identified earlier as the positive control (unpublished data).

Phylogenetic analyses based on 16S rDNA including defined type strains of different *Bradyrhizobium*, *Ensifer* and *Microvirga* spp. showed that in all three genera the isolated strains from Jizan soil represent undefined species (Fig. 3). *Microvirga* was only recently discovered as a nitrogen-fixing symbiotic genus, and *Microvirga* sp. SA342 is phylogenetically related to the symbiotic species *M. subterranea* (Kanso and Patel, 2003; Radl et al., 2014) (Fig. 3a). *Ensifer* sp. SA331 showed to be most closely related to *E. kostiensis* (also known as *Ensifer/Sinorhizobium kostiense*), a species known to nodulate several legume species in Asia and Africa (Nick et al., 1999; Räsänen et al., 2001; Merabet et al., 2010; Sankhla et al., 2017) (Fig. 3b). In case of *Bradyrhizobium*, the newly isolated strains from Jizan soil classify in two groups. One group, represented by a single strain, SA281, is closely related to *B. elkanii* and *B. pachyrhizi*, whereas the other 11 strains form a separate branch that is remotely related to *B. neotropica* (Fig. 3c). This latter species was only recently identified in root nodules of *Centrolobium paraense* originating from the Brazilian amazon (Zilli et al., 2014). To classify the isolated rhizobium strains, BOX PCR-based genetic fingerprinting was conducted. This identified 29 groups with a unique DNA fingerprint; 12 *Bradyrhizobium* groups, 8 *Microvirga* groups, and 9 groups representing *Ensifer* spp.. For the 12 *Bradyrhizobium* groups, a representative strain was randomly selected from each group, and for *Microvirga* and *Ensifer*, 1 strain was randomly selected as a representative from each genus (Table 1).

Taken together, this experiment revealed that *I. argentea* grown in Jizan soil can be nodulated by rhizobium species of the genera *Microvirga*, *Ensifer*, and *Bradyrhizobium*, of which the latter is most genetically diverse. This led us to conclude that *I. argentea* is promiscuous for rhizobia from at least three unrelated genera.

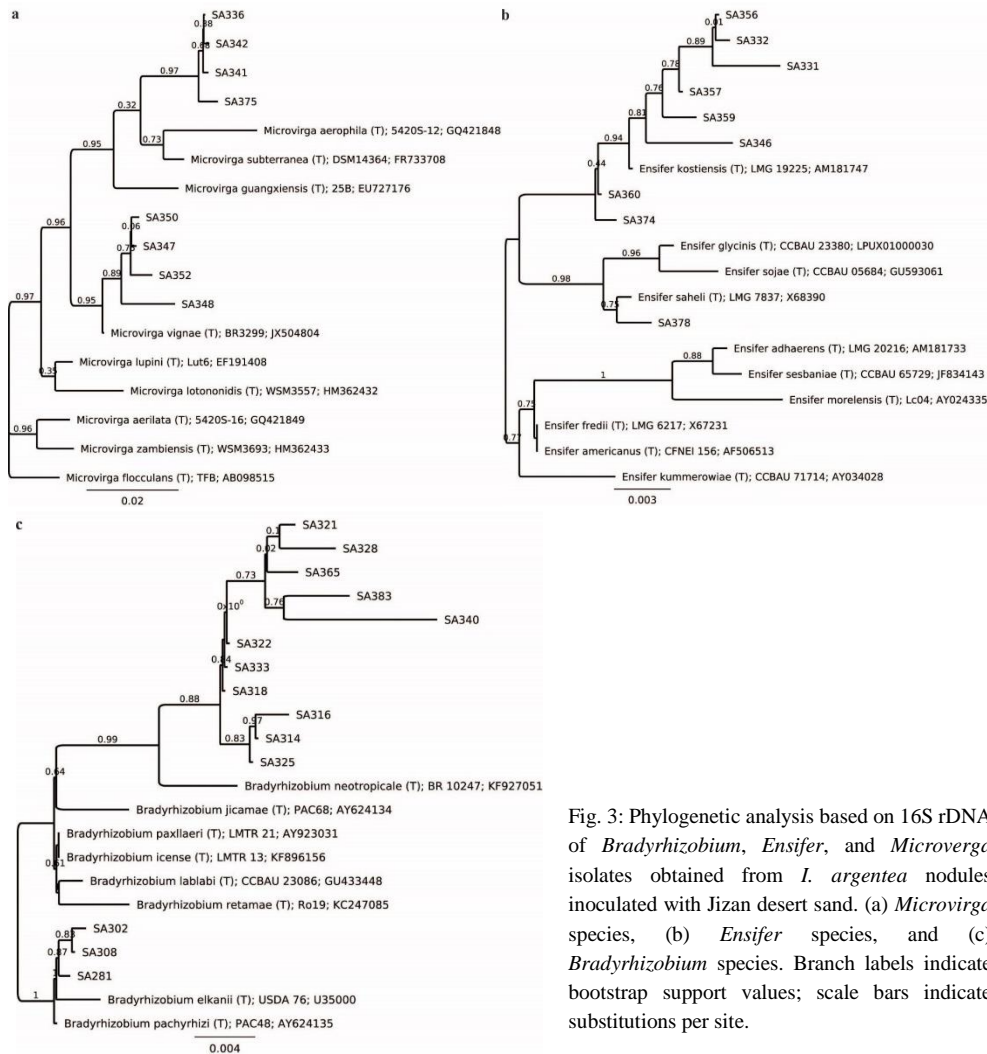


Fig. 3: Phylogenetic analysis based on 16S rDNA of *Bradyrhizobium*, *Ensifer*, and *Microverga* isolates obtained from *I. argentea* nodules inoculated with Jizan desert sand. (a) *Microverga* species, (b) *Ensifer* species, and (c) *Bradyrhizobium* species. Branch labels indicate bootstrap support values; scale bars indicate substitutions per site.

Bradyrhizobium* sp. strain SA340 is a highly infective microsymbiont for *I. argentea

Nodulation studies using *Bradyrhizobium* sp. strain SA383 as inoculum showed to be very efficient. Here we aimed to determine the infectiveness of the different strains isolated from soil. To do so, we first developed a protocol that will allow BOX-PCR based fingerprinting of rhizobia inside nodules. The fingerprint pattern of cultured rhizobia and infected root nodules was compared (unpublished data). This revealed that the fingerprint patterns of rhizobia and *I. argentea* are very distinct. This allows identification of the most dominant rhizobium strain inside a root nodule.

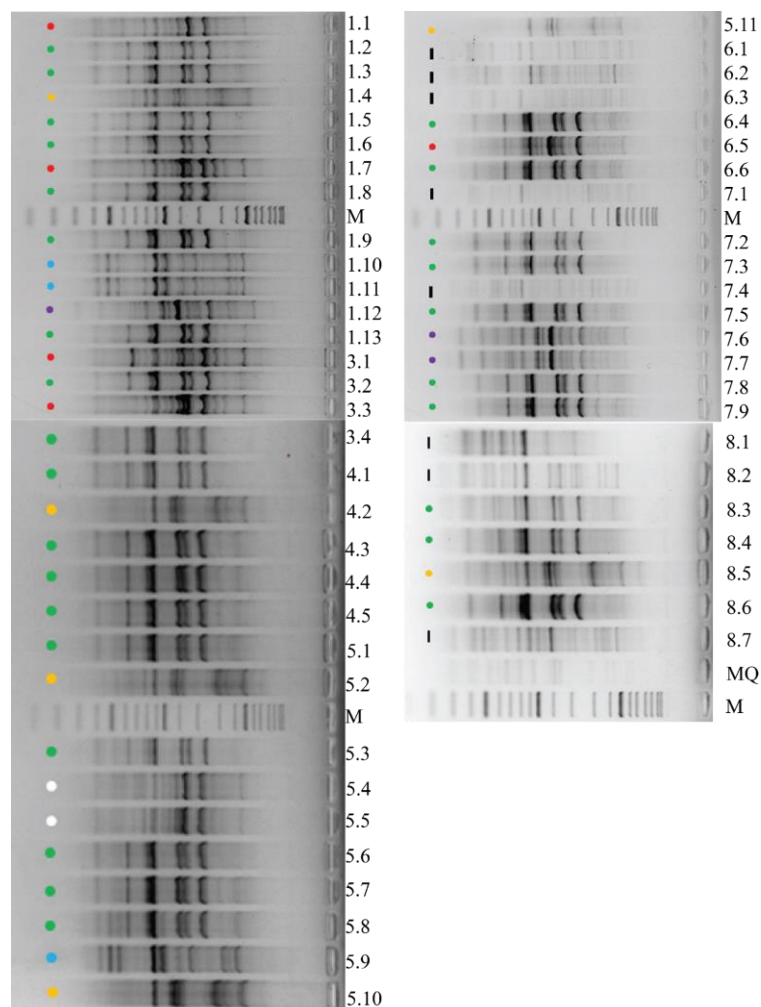


Fig. 4: BOX-PCR based DNA fingerprinting pattern of isolated bacteria by BOX-PCR. Green, SA340 (*Bradyrhizobium*); Orange, SA333 (*Bradyrhizobium*); Red, SA314 (*Bradyrhizobium*); Purple, SA331 (*Ensifer*); Blue, SA342 (*Microvirga*); White, SA328 (*Bradyrhizobium*); 8 Black stripes, too weak for identification. The codes above the gel indicate from which plant (replicate plants 1, 3, 4, 5, 6, 7 & 8) and the numbers after the dot indicate which nodule arranged from smallest to largest.

I. argentea plants were inoculated with a synthetic community of 17 rhizobial strains; the 14 different fingerprint groups identified in this study (Table 1), supplemented with 3 strains representing distinct groups identified earlier (2 *Rhizobium* and 1 *Microvirga*, unpublished data). In total 55 nodules from 7 plants were analyzed 4 weeks post inoculation. BOX PCR-based fingerprinting of nodules was identified. For 47 nodules (out of 55) the DNA

fingerprint could be matched with a BOX-PCR pattern of one of 17 the bacteria used as inoculum, whereas in 8 cases the signal was too weak for identification (Fig. 4). Interestingly, the 47 nodules that could be genotyped were infected with rhizobia of only 6 groups; 4x *Bradyrhizobium* spp., 1x *Ensifer* spp., and 1x *Microvirga* spp.. Of these *Bradyrhizobium* sp. strain SA340 showed to be most infective, as it colonized 28 (out of 55) nodules (Table 2). This suggests that despite its promiscuous nature, *I. argentea* nodulation efficiency varies between compatible strains.

To determine whether *Bradyrhizobium* sp. strain SA340 not only is a very infective strain on *I. argentea*, but also symbiotically effective, we compared its plant growth promoting effect with representative *Ensifer* and *Micorvirga* strains SA331 and SA342. Four weeks post inoculation with single strains, the root and shoot dry weights of the *I. argentea* plantlets was determined (Fig. 5). This showed that all 3 strains improved plant growth, with *Bradyrhizobium* sp. strain SA340 as most effective strain. This led us to conclude that this strain not only is highly infective, but also effective with respect of plant growth stimulation.

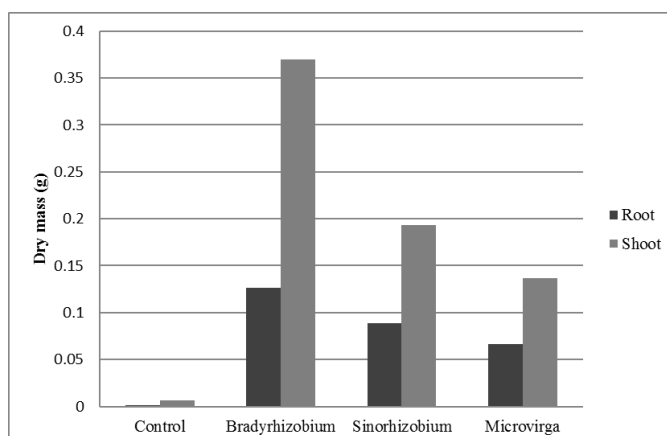


Fig. 5: Dry mass of root and shoot of *I. argentea* inoculated with *B. genosp* SA340, *E. kostiense* SA331, and *Microverga* sp. SA342, compared to uninoculated control plants. Dry mass of root and shoot were mean dry weight of 6 plants, 4 weeks post inoculation.

Discussion

We demonstrated that *I. argentea* is highly promiscuous and can form nodules with a diverse range of rhizobia, including different *Bradyrhizobium*, *Ensifer*, and *Microvirga* species. For the latter this is the first report that a bacterial species of this genus can nodulate *Indigofera* spp.. By using BOX PCR fingerprinting on root nodules, we demonstrated that a single *Bradyrhizobium* spp. represented by strain SA340, is most infective in competition studies. Additionally, we demonstrate that this strain has profound growth promoting effects on *I. argentea* when grown under low nitrogen regime. Taken together, we concluded that despite the promiscuous nature of *I. argentea*, it is predominantly nodulated by specific strains which can have a high symbiotic effectiveness.

Material and Methods

Plant material, seeds germination and root inoculation

Nodules and seeds of the pioneer legume *I. argentea* and sand were collected from Jizan desert, Saudi Arabia. Seeds germination was conducted as described in Chapter 2. Sands collected from Jizan desert was used as “inoculant” on the seedlings of the second *I. argentea* generation in the greenhouse. The nodules that were harvested 4 weeks after inoculation in the greenhouse and collected from Jizan desert were used for rhizobial strains isolation.

Rhizobial strains isolation

The nodules were collected from plants collected in the desert as well as grown in sand and clay mixed with 10 g sand from desert in the greenhouse. Nodules were surface sterilized with 96% ethanol for 30 sec, followed by 4% sodium hypochlorite (commercial bleach) for 60 sec. Surface sterilized nodules were squashed in 200µl 0.9% NaCl solution. Diluted suspension was plated on yeast mannitol agar plates (YEM). Plates were incubated at 28°C for 1-7 days. Isolated rhizobia were characterized by sequencing 16S rDNA using the primer set: 63F-5'-CAGGCCTAACACATGCAAGTC-3' and 1389R-5'-ACGGGCGGTGTGTACAAG-3' (Muresu et al., 2008). The isolated rhizobia were listed in Table 1.

Phylogenetic Analysis

Phylogenetic analyses were performed based on trimmed suspicious nucleotides from the relatively short 16S sequences of isolated strains in Table 1, in combination with 16S sequences of type strains retrieved from the RDP database (see Fig. 4). Alignment was done with MAFFT 7.017 and unrooted phylogenetic trees were produced with FastTree 2.1.5.

Genetic variety analysis of rhizobia strains by using BOX-PCR fingerprints

Rhizobia strains were selected based on genetic variation of all isolates. PCR amplification of repetitive regions of the genomic DNA (rep-PCR) was performed with BOX-A1R primer (5'-CTACGGCAAGGCGACGCTGACG-3'), which can result in high resolution fingerprints (Lanoot et al., 2004; Rademaker et al., 2004). And the PCR program used was: 95°C for 2 min followed by 35 cycles each of 94°C for 3s, 92°C for 30s, 50°C for 1 min and 72°C for 8 min, followed by 72°C for 8 min and then cooling to 8°C. PCR products were visualized at 80 voltage for 2.5 hours on 2% agarose gel. Rhizobia strains with the same fingerprint were grouped in one box. One representative strain was selected per BOX group.

Identification of the most competitive rhizobia strain in consortium from nodules

Nodulation assay was performed by growing *I. argentea* plants on sterilized river sands supplemented with rhizobia consortium. Selected rhizobia strains were cultivated in 10 ml

YEM liquid medium at 28 °C and shaken at 220 rpm. Each of bacterial cultures was washed three times with 0.9% NaCl and set to OD600 = 0.001 with Färhaues solution. The rhizobia consortium was made by mixing all 17 selected rhizobia strains identically. Sterilized river sand was inoculated with rhizobia consortium and homogenized prior to transplanting *I. argentea* seedlings. After transplanting two *I. argentea* seedlings per pot, another 1 ml consortium culture was added to root of each seedlings. Then plants were cultivated in WEISS KAST (35°C/25°C day/night temperatures; 250 µmol light m⁻² s⁻¹ at plant level during 12 h/d; 70% relative humidity).

Nodules were collected 6 weeks after transplanting. These nodules were surface sterilized by soaking in 96% ethanol for 20 seconds, washing with sterilized MQ water for 3 times and then putting in 4% bleach for maximum 3 minutes, wash off bleach by sterilized MQ water for 6 times. Nodules were crushed with a metal bead. Put each nodule in a 2 ml tube with a sterile metal bead. Put tubes in liquid nitrogen and then and crush the nodules two times with a tissue lyser (50 times/second for 30 seconds). Suspending crushed nodules in 200 µl sterile MQ water, this suspension was further diluted for 10 times of which 1 µl was used as BOX-PCR DNA template. PCR products were visualized on 2% agarose gel and fingerprint obtained from each nodules were compared to each of 17 rhizobia strains.

Tissue embedding, sectioning, section staining and microscopy

The longitudinal section and microscopy were conducted as described in Chapter 4.

Chapter 3

Table 1: *I. argentea* micorsymbionts isolated from Jizan dessert soil catogerized according to BOX-PCR fingerprinting. Representative strain of each group used for the competition studies marked in bold.

Strain	Genus	nifH gene
SA281	<i>Bradyrhizobium elkanii</i>	
SA302	<i>Bradyrhizobium elkanii</i>	
SA308	<i>Bradyrhizobium elkanii</i>	
SA314	<i>Bradyrhizobium genosp.</i>	
SA315	<i>Bradyrhizobium genosp.</i>	
SA324	<i>Bradyrhizobium genosp.</i>	
SA337	<i>Bradyrhizobium genosp.</i>	
SA338	<i>Bradyrhizobium genosp.</i>	
SA339	<i>Bradyrhizobium genosp.</i>	
SA361	<i>Bradyrhizobium genosp.</i>	
SA317	<i>Bradyrhizobium genosp.</i>	
SA318	<i>Bradyrhizobium genosp.</i>	
SA320	<i>Bradyrhizobium genosp.</i>	
SA323	<i>Bradyrhizobium genosp.</i>	
SA321	<i>Bradyrhizobium genosp.</i>	
SA335	<i>Bradyrhizobium genosp.</i>	
SA343	<i>Bradyrhizobium genosp.</i>	
SA345	<i>Bradyrhizobium genosp.</i>	
SA322	<i>Bradyrhizobium genosp.</i>	
SA326	<i>Bradyrhizobium genosp.</i>	
SA330	<i>Bradyrhizobium genosp.</i>	
SA325	<i>Bradyrhizobium genosp.</i>	
SA319	<i>Bradyrhizobium genosp.</i>	
SA328	<i>Bradyrhizobium genosp.</i>	
SA334	<i>Bradyrhizobium genosp.</i>	
SA327	<i>Bradyrhizobium genosp.</i>	
SA329	<i>Bradyrhizobium genosp.</i>	
SA333	<i>Bradyrhizobium genosp.</i>	
SA365	<i>Bradyrhizobium genosp.</i>	
SA369	<i>Bradyrhizobium genosp.</i>	
SA367	<i>Bradyrhizobium genosp.</i>	
SA368	<i>Bradyrhizobium genosp.</i>	
SA379	<i>Bradyrhizobium genosp.</i>	
SA383	<i>Bradyrhizobium genosp.</i>	
SA316	<i>Bradyrhizobium genosp.</i>	
SA373	<i>Bradyrhizobium genosp.</i>	
SA340	<i>Bradyrhizobium genosp.</i>	
SA355	<i>Bradyrhizobium genosp.</i>	
SA360	<i>Ensifer sp. NFY8</i>	
SA331	<i>Ensifer kostiense</i>	
SA332	<i>Ensifer kostiense</i>	
SA346	<i>Ensifer kostiense</i>	

Variation in infectiveness of rhizobium strains

SA356	<i>Ensifer kostiense</i>	
SA357	<i>Ensifer kostiense</i>	
SA359	<i>Ensifer kostiense</i>	
SA374	<i>Ensifer sp. NFY8</i>	
SA378	<i>Ensifer sp. Mh9</i>	
SA336	<i>Microvirga sp. S-MI1b</i>	<i>Mesorhizobium temperatum,</i>
SA341	<i>Microvirga sp. S-MI1b</i>	<i>Mesorhizobium temperatum,</i>
SA342	<i>Microvirga sp. S-MI1b</i>	<i>Mesorhizobium temperatum,</i>
SA347	<i>Microvirga vignae</i>	<i>Mesorhizobium temperatum,</i>
SA348	<i>Microvirga vignae</i>	<i>Mesorhizobium temperatum,</i>
SA350	<i>Microvirga vignae</i>	<i>Mesorhizobium temperatum,</i>
SA352	<i>Microvirga vignae</i>	<i>Mesorhizobium temperatum,</i>
SA375	<i>Microvirga sp. S-MI1b</i>	<i>Mesorhizobium temperatum,</i>

Table 2: List of strains were identified from 55 nodules of *I. argentea*.

Strains	# of nodules	BOX PCR pattern	Ratio (%)
SA340	28	Green	50.9
SA333	6	Orange	10.9
SA314	5	Red	9.1
SA331	3	Purple	5.5
SA342	3	Blue	5.5
SA328	2	White	3.6
Others*	8	Black	14.5

* There were 8 strains identified from 8 nodules. The BOX-PCR patterns of these strains are hardly grouped.

References

- Andrews, M., and Andrews, M.E. (2017). Specificity in Legume-Rhizobia Symbioses. *Int. J. Mol. Sci.* 18: 705.
- Angus, A.A., Lee, A., Lum, M.R., Shehayeb, M., Hessabi, R., Fujishige, N.A., Yerrapragada, S., Kano, S., Song, N., and Yang, P. (2013). Nodulation and effective nitrogen fixation of *Macroptilium atropurpureum* (siratro) by *Burkholderia tuberum*, a nodulating and plant growth promoting beta-proteobacterium, are influenced by environmental factors. *Plant Soil* 369:543-562.
- Clark, V.R., Schrir, B.D., and Barker, N.P. (2015). Two new species of *Indigofera* L. (Leguminosae) from the Sneeuberg Centre of Floristic Endemism, Great Escarpment (Eastern and Western Cape, South Africa). *PhytoKeys* 48: 29-41.
- Dresler-Nurmi, A., Fewer, D.P., Räsänen, L.A., and Lindström, K. (2009). The diversity and evolution of *Rhizobia*. *Microbiol Monogr* 8: 3-41.
- Gyaneshwar, P., Hirsch, A.M., Moulin, L., Chen, W.M., Elliott, G.N., Bontemps, C., los Santos, P.E., Gross, E., dos Reis, F.B., Jr., Sprent, J.I., Young, J.P.W., and James, E.K. (2011). Legume-nodulating betaproteobacteria: diversity, host range and future prospects. *Molecular Plant-Microbe Interactions* 24: 1276-1288.
- Geurts, R., Heidstra, R., Hadri, A.E., Downie, J.A., Franssen, H., Van Kammen, A., and Bisseling, T. (1997). Sym2 of *Pisum sativum* is involved in a Nodulation factor perception mechanism that controls the infection process in the epidermis. *Plant Physiol* 115: 351-359.
- 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 calcium spiking. *Plant J* 77: 146-159.
- Jiao, Y.S., Liu, Y.H., Yan, H., Wang, E.T., Tian, C.F., Chen, W.X., Guo, B.L., and Chen, W.F. (2015). Rhizobial diversity and nodulation characteristics of the extremely promiscuous legume *Sophora flavescens*. *Molecular Plant-Microbe Interactions* 28(12): 1338-1352.
- Jordan, D.C. (1982). Transfer of *Rhizobium japonicum* Buchanan 1980 to *Bradyrhizobium* gen. nov., a genus of slow-growing, root nodule bacteria from leguminous plants. *Int J Syst Bacteriol*, 32: 136-139.
- Kanso, S., and Patel, B.K. (2003). *Microvirga subterranea* gen. nov., sp. nov., a moderate thermophile from a deep subsurface Australian thermal aquifer. *International Journal of Systematic and Evolutionary Microbiology* 53(2): 401-406.
- Kawaharada, Y., Kelly, S., Nielsen, M.W., Hjuler, C.T., Gysel, K., Muszyński, A., Carlson, R.W., Thygesen, M.B., Sandal, N., Asmussen, M.H., Vinther, M., Andersen, S.U., Krusell, L., Thirup, S., Jensen, K.J., Ronson, C.W., Blaise, M., Radutoiu, and S., Stougaard, J. (2015). Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature* 523:308-312.
- Kawaharada, Y., James, E.K., Kelly, S., Sandal, N., and Stougaard, J. (2017). The Ethylene Responsive Factor Required for Nodulation 1 (ERN1) transcription factor is required for infection-thread formation in *Lotus japonicus*. *Mol Plant Microbe Interact* 30: 194-204.
- Kuykendall, L.D., Saxena, B., Devine, T.E., and Udell, S.E. (1992). Genetic diversity in *Bradyrhizobium japonicum* Jordan 1982 and a proposal for *Bradyrhizobium elkanii* sp. nov. *Can J Microbiol* 38: 501-505.
- Lanoot, B., Vancanneyt, M., Dawyndt, P., Cnockaert, M., Zhang, J.L., Huang, Y., Liu, Z.H., and Swings, J. (2004). BOX-PCR fingerprinting as a powerful tool to reveal synonymous names in the genus *Streptomyces*. Emended descriptions are proposed for the species *Streptomyces cinereorectus*, *S. fradiae*, *S. tricolor*, *S. colombiensis*, *S. filamentosus*, *S. vinaceus* and *S. phaeopurpureus*. *Systematic and Applied Microbiology* 27(1): 84-92.
- Leelahawong, C., Nuntagij, A., Teaumroong, N., Boonkerd, N., and Pongsilp, N. (2010). Characterization of root-nodule bacteria isolated from the medicinal legume *Indigofera tinctorial*. *Ann Microbiol* 60: 65-74.
- Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J.C., and Dénarié, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* 344(6268): 781-784.
- Lira, M.A., Nascimento, L.R.S., and Fracetto, G.G.M. (2015). Legume-rhizobia signal exchange: promiscuity and environmental effects. *Frontiers in Microbiology* 6: 945.
- Merabet, C., Martens, M., Mahdhi, M., Zakhia, F., Sy, A., Le Roux, C., Domergue, O., Coopman, R., Bekki, A., Mars, M., Willems, A., and De Lajudie, P. (2010). Multilocus sequence analysis of root nodule isolates from *Lotus arabicus* (Senegal), *Lotus creticus*, *Argyrolobium uniflorum* and *Medicago sativa* (Tunisia) and description of *Ensifer numidicus* sp. nov. and *Ensifer garamanticus* sp. nov. *Int J Syst Evol Microbiol* 60: 664-674.

- Michael, J.T., Celia, M., and Paul, A.H. (1991). Formation and structure of root nodules induced on *Macroptilium atropurpureum* inoculated with various species of *Rhizobium*. Canadian Journal of Botany 69(7): 1520-1532.
- Minamisawa, K., Onodera, S., Tanimura, Y., Kobayashi, N., Yuhashi, K., and Kubota, M. (1997). Preferential nodulation of *Glycine max*, *Glycine soja* and *Macroptilium atropurpureum* by two *Bradyrhizobium* species *japonicum* and *elkanii*. FEMS Microbiol. Ecol. 24: 49–56.
- Moulin, L., Munive, A., Dreyfus, B., and Boivin-Masson, C. (2001). Nodulation of legumes by members of the beta-subclass of Proteobacteria. Nature 411(6840): 948-950.
- Moulin, L., James, E.K., Klonowska, A., Faria, S.M., and Simon, M.F. (2015). Phylogeny, diversity, geographical distribution, and host range of legume-nodulating Betaproteobacteria: what is the role of plant taxonomy? In: de Bruijn FJ, ed. Biological nitrogen fixation. Chichester, UK: John Wiley & Sons Inc, 177-190.
- Muresu, R., Polone, E., and Sulas, L. (2008). Coexistence of predominantly nonculturable rhizobia with diverse, endophytic bacterial taxa within nodules of wild legumes. FEMS Microbiol Ecol 63:383-400.
- Nick, G., De Lajudie, P., Eardly, B.D., Suomalainen, S., Paulin, L., Zhang, X.P., Gillis, M., and Lindström, K. (1999). *Sinorhizobium arboris* sp. nov. and *Sinorhizobium kostense* sp. nov., isolated from leguminous trees in Sudan and Kenya. Int. J. Syst. Bacteriol. 49: 1359-1368.
- Op Den Camp, R., Streng, A., and De Mita, S. (2011). LysM-type mycorrhizal receptor recruited for rhizobium symbiosis in nonlegume *Parasponia*. Science 331: 909-912.
- Perret, X., Staehelin, C., and Broughton, W.J. (2000). Molecular basis of symbiotic promiscuity. Microbiol. Mol. Biol. Rev 64:180–201.
- Rademaker, J.L., Louws, R.J., Versalovic, J., and Bruijn, F.J. (2004). Characterization of the diversity of ecologically important microbes by rep-PCR genomic fingerprinting. In Kowalchuk, G.A. et al. (Eds.) Molecular microbial ecology manual (pp. 611-643). Vol. 1. Boston, MA: Kluwer Academic.
- Räsänen, L.A., Elvang, A.M., Jansson, J., and Lindström, K. (2001). Effect of heat stress on cell activity and cell morphology of the tropical rhizobium, *Sinorhizobium arboris*. FEMS Microbiol Ecol 34:267-278.
- Radl, V., Simões-Araújo, J.L., Leite, J., Passos, S.R., Martins, L.M., Xavier, G.R., Rumjanek, N.G., Baldani, J.I., and Zilli, J.E. (2014). *Microvirga vignae* sp. nov., a root nodule symbiotic bacterium isolated from cowpea grown in semi-arid Brazil. International journal of systematic and evolutionary microbiology. 64(3): 725-730.
- Radutoiu, S., Madsen, L.H., Madsen, E.B., Jurkiewicz, A., Fukai, E., Quistgaard, E.M.H., Albrektsen, A.S., James, E.K., Thirup, S., and Stougaard, J. (2007). LysM domains mediate lipochitin-oligosaccharide recognition and NFR genes extend the symbiotic host range. EMBO J 26: 3923-3935.
- Sankhla, I.S., Tak, N., Meghwal, R.R., Choudhary, S., Tak, A., Rathi, S., Sprent, J.I., James, E.K., and Gehlot, H.S. (2017). Molecular characterization of nitrogen fixing microsymbionts from root nodules of *Vachellia (Acacia) jacquemontii*, a native legume from the Thar Desert of India. Plant Soil 410: 21-40.
- Schrire, B.D. (2005). Tribe Phaseoleae. Pp. 361-364. In: Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (eds.), Legumes of the World. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Schrire, B.D., Lavin, M., Barker, N.P., and Forest, F. (2009). Phylogeny of the tribe Indigoferae (Leguminosae-Papilionoideae): Geographically structured more in succulent-rich and temperate settings than in grass-rich environments. Am J Bot 96(4): 816-852.
- Sprent, J.I., Ardley, J., and James, E.K., (2017). Biogeography of nodulated legumes and their nitrogen fixing symbionts. New Phytologist 215: 40-56.
- Vinuesa, P., Rojas-Jiménez, K., Contreras-Moreira, B., Mahna, S.K., Prasad, B.N., and Moe, H. (2008). Multilocus sequence analysis for assessment of the biogeography and evolutionary genetics of four *Bradyrhizobium* species that nodulate soybeans on the Asiatic continent. Appl. Environ. Microbiol. 74: 6987-6996.
- Xu, L.M., Ge, C., Cui, Z., Li, J., and Fan, H. (1995). *Bradyrhizobium liaoningense* sp. nov., isolated from the root nodules of soybeans. International Journal of Systematic Bacteriology. 45(4): 706-711.
- Yang, S., Tang, F., Gao, M., Krishnan, H.B., and Zhu, H. (2010). R gene-controlled host specificity in the legume-rhizobia symbiosis. Proc. Natl. Acad. Sci. USA 107: 18735-18740.
- Zilli, J.E., Barauna, A.C., Da Silva, K., De Meyer, S.E., Farias, E.N.C., Kaminski, P.E., Da Costa, I.B., Ardley, J.K., Willems, A., Camacho, N.N., Dourado, F.d.S., and O'Hara, G. (2014). *Bradyrhizobium neotropicae* sp. nov., isolated from effective nodules of *Centrolobium paraense*. International Journal of Systematic and Evolutionary Microbiology 64(12): 3950-3957.

Chapter 4

Evolution of determinate and indeterminate nodules in Papilionoideae

Guiling Ren¹, Olga Kulikova¹, Arjan van Zeijl¹, Hukam S. Gehlot², Carolien Franken¹, Robin van Velzen¹, Tingting Xiao¹, Henk Franssen¹, and Ton Bisseling¹

¹ Laboratory of Molecular Biology, Plant Sciences Group, Wageningen University and Research Centre, Wageningen, The Netherlands

² BNF & Stress Biology laboratory, Department of Botany, J.N.Vyas University, Jodhpur, Rajasthan 342001, India

Abstract

- The infection and morphogenesis process of nodules of *Indigofera* and *Tephrosia* species were described in detail in this chapter.
- In both *Indigofera* and *Tephrosia* species, primordia are formed from the outermost cortical cells and the meristematic region disappears at an early stage of development. These characteristics resemble the determinate nodule type. Strikingly, at later stages clusters of dividing infected cells are present at the apex. Using ENOD2 as a molecular marker we showed that fully differentiated nodule parenchyma can form such clusters of dividing cells.
- We propose that the common ancestor of the Papilionoideae subfamily made determinate nodules and indeterminate nodule type first evolved in the IRLC and robinoids clade.
- The evolution of determinate and indeterminate nodule types was discussed.

Key words: determinate nodule, indeterminate nodule, nodule evolution, cluster of dividing cells, *Indigofera*, *Tephrosia*.

Introduction

Legumes can establish a nitrogen-fixing nodulation symbiosis with phylogenetically diverse bacteria (Remigi et al., 2016) that are collectively named rhizobia. The developmentally and biochemically diverse nodules formed by legumes are typically classified as being either determinate or indeterminate (Hirsch, 1992; Sprent, 2001). A major difference between these two types is the life span of their meristem. Indeterminate nodules have a persistent meristem at their apex and attain an elongated shape. In contrast, determinate nodules have a meristem at the periphery and this is only active at early stages of development (Bergersen, 1958). Typical determinate nodules have a spherical shape. The best characterised indeterminate nodules are those of the IRLC (Inverted Repeat Lacking Clade: e.g. *Medicago*, clover, pea, Fig. 1) and typical determinate nodules are formed by genera in a branch of the millettoids clade (e.g. *Glycine*, *Vigna* and *Phaseolus*), although distinctive determinate nodules occur elsewhere in the family (Lavin et al., 2001). When we describe (in)determinate nodule types in this chapter, we refer to these two clades.

The millettoids, *Indigofera*, robinoids, and IRLC groups (hereafter named ‘selected groups’) belong to the Papilionoideae subfamily and contain most legume crops, as well as the model legumes *Lotus japonicus* (*Lotus*) and *Medicago truncatula* (*Medicago*). The evolutionary relationships of 24 genera belonging to these 4 groups as well as our current knowledge concerning the nodule type that they form is shown in Fig.1 (modified from Oono et al., 2010). As genera comprising mirbelioids and genistoids have been reported to form indeterminate nodules, the current published data suggest that the common ancestor of the 4 selected groups formed indeterminate nodules (Fig. 1, node 15). This is further supported by the fact that the first split in the dalbergioids clade (Fig. 1, node 1) divides the group into genera reported to form indeterminate nodules (tribe Amorpheae: e.g. *Amorpha* and *Dalea*) and a group with a unique type of nodule (‘aeschynomenoid’; Lavin et al., 2001) that is likely to have originated from determinate nodule (Doyle, 2011).

If the common ancestor of the 4 selected groups indeed formed indeterminate nodules, the determinate nodule type seems to have evolved twice in the millettoids clade (nodes 24 and 25) and once in the robinoids clade (node 26).

Indigofera is sister to the millettoids clade (Fig. 1), and so *Indigofera*, by far the largest genus within the tribe Indigofereae (Schrire et al., 2009), is a key genus in understanding the evolution of determinate and indeterminate nodule types. *Indigofera* nodules have been shown to have a cluster of dividing cells at their apex and therefore were classified as indeterminate nodules (Oono et al., 2010; Lemaire et al., 2015). However, although *Indigofera* nodules have this apical cluster of dividing cells, they also have some characteristics, for example lenticels (Gehlot et al., 2012), that are common in determinate

nodules (Pankhurst and Sprent, 1975; Sprent, 2001). As *Indigofera* has a crucial phylogenetic position (Fig. 1), we reinvestigated whether it indeed forms genuine indeterminate nodules.

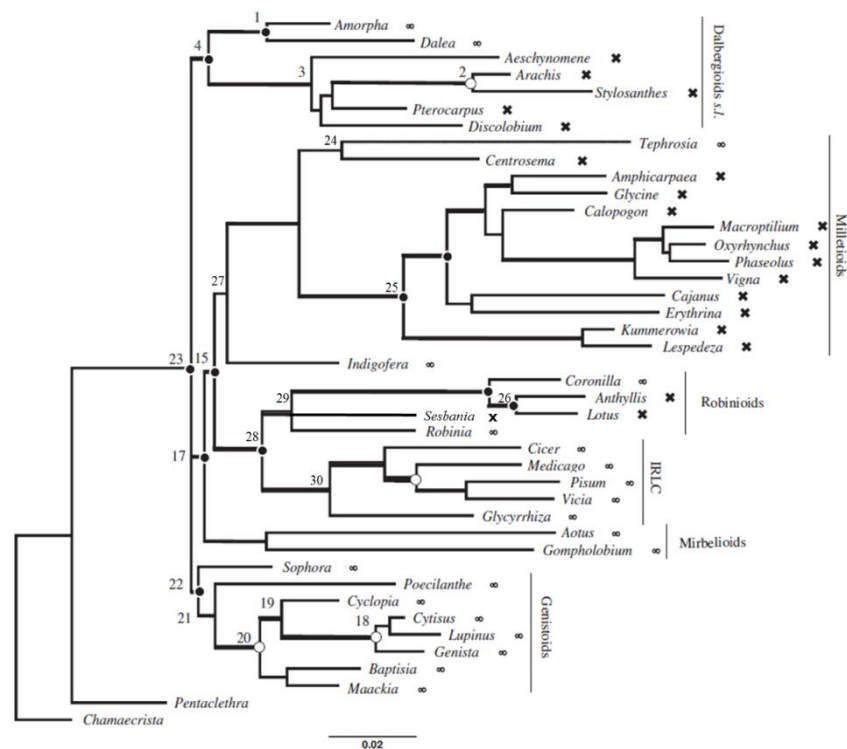


Fig.1 Distribution of indeterminate and determinate nodule types within the Papilionoideae subfamily according to Oono et al. (2010). Thickened branches indicate support of posterior probability ≥ 0.95 . Character for indeterminate (infinity) and determinate nodule type (crosses) is indicated next to the genus name.

Due to the difference in meristem persistence there are some additional characteristics that distinguish the two nodule types. Indeterminate nodules have a meristem at the apex and this meristem adds cells in proximal direction to the different nodule tissues. As a consequence, mature tissues are absent at the nodule apex. This is for example the case for the peripheral tissues like endodermis and nodule parenchyma. In contrast, in determinate nodules these tissues completely surround the determinate nodule and the vascular bundles, located in the nodule parenchyma, even fuse. Mature nodule endodermis cells form Casparian strips which can be used to detect them (Brundrett et al., 1988). The *ENOD2* gene is specifically expressed in the nodule parenchyma in legumes belonging to the IRLC (pea, *Medicago*, *Vicia*) (Van de Wiel et al., 1990a; Van de Wiel et al., 1990b; Vijn et al., 1995), robinioideae clade (*Sesbania*, *Lotus*) (Lauridsen et al., 1993; Goormachtig et al., 1998) and millettioideae clade (*Glycine*) (Van

de Wiel et al., 1990a). Therefore, it is probable that *Indigofera* species have an orthologue with a similar expression pattern, which then could be used to study the development of the nodule parenchyma. Although the histological studies have only been done on a few species (e.g. pea, Medicago, clover, soybean), it is probably that the characteristics of the peripheral tissues are generic for nodules lacking an apical meristem that adds cells to the different nodule tissues.

Based on the best studied nodule developmental processes the two nodule types seem to be formed from different root cell layers. Indeterminate nodules of the IRLC tribe (e.g. Medicago) are formed from inner cortical layers (Xiao et al., 2014). In contrast, determinate nodules of the milletioids clade (e.g. *Glycine*) are formed from outer cortical layers (Turgeon and Bauer, 1982) and those from the robinoids tribe (e.g. Lotus, *Sesbania*) from the middle cortical layers (Ndoye et al., 1994; Madsen et al., 2010). Further, determinate nodules almost completely lack vacuoles in their infected cells, whereas infected cells of indeterminate nodules have a clear vacuole (Sprent, 2001).

We reanalysed *Indigofera* nodule development and this showed that it has no persistent apical meristem. The central tissue of rather young nodules becomes completely surrounded by endodermis and nodule parenchyma and at this stage an apical meristem is lacking. Therefore, we consider it to be a determinate nodule. The cluster of dividing cells often present at the apex of older nodule is formed secondarily from fully differentiated cells. The finding that *Indigofera* forms determinate nodules prompted us to reinvestigate *Tephrosia* nodule development as it remained the only genus within the milletioids tribe known to form indeterminate nodules (Fig. 1) (Oono et al., 2010). *Tephrosia* was proposed to form indeterminate nodules based on the presence of dividing cells (Gehlot et al., 2012). We showed that its development is very similar to that of *Indigofera* in that it forms determinate nodules and clusters of dividing cells can be formed secondarily. Our findings strongly suggest that the common ancestor of the milletioids and *Indigofera* formed determinate nodules. We discuss the implication of this hypothesis for the evolution of the two nodule types.

Results

Cluster of dividing cells at the apex of *I. argentea* nodules

To characterize the cytology of elongated *I. argentea* nodules, seedlings (8-days-old) were inoculated with *Bradyrhizobium elkanii* SA281. Nodules were harvested 4 weeks after inoculation and elongated nodules were sectioned (Fig. 2a). This showed that a cluster of small dividing cells is present at the apex (Fig. 2a, b), as was reported previously (Gehlot et al., 2012). Further, these nodules contain characteristic tissues, including the central tissue, which is surrounded by the nodule parenchyma, endodermis and nodule cortex. *I. argentea*

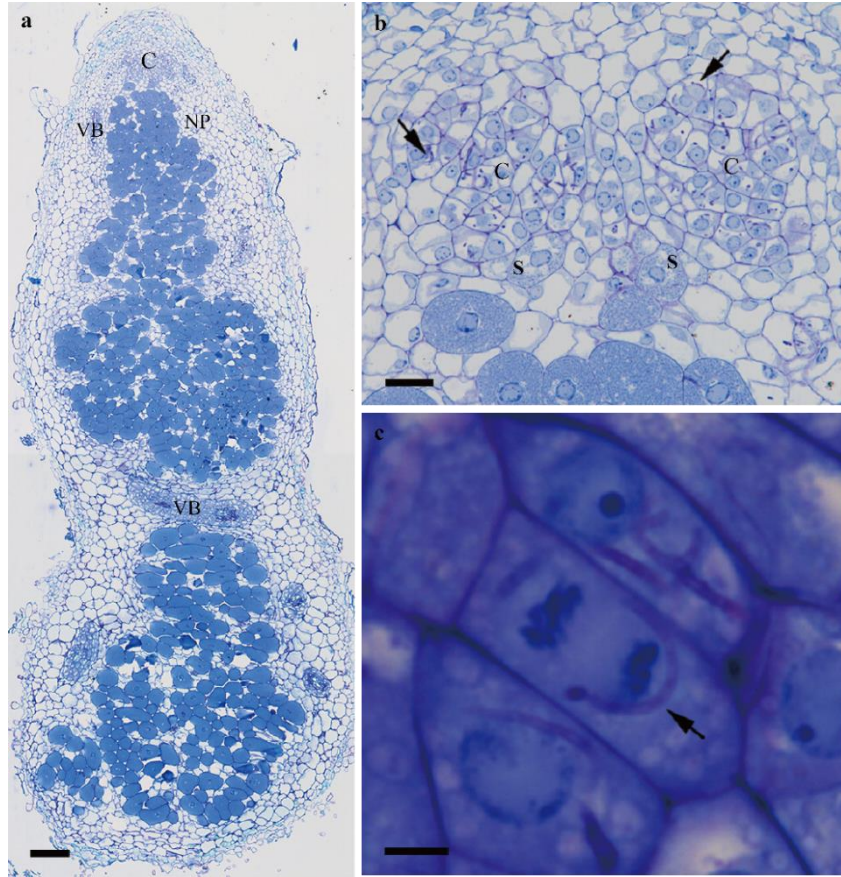


Fig. 2 Longitudinal sections of *I. argentea* stained with toluidine blue O. (a) Section of an elongated nodule, cluster of dividing cells is present at the apex; (b) Infection threads are present in the vast majority of these cells even in the most distal ones (arrow), bacteria are released in some cells (arrowhead); (c) Dividing cells can contain infection threads (arrow). C, cluster of dividing cells; VB, vascular bundle; NP, nodule parenchyma; S: symbiosome; Scale bars: a, 150 μ m; b, 25 μ m; c, 5 μ m.

nodules have been classified as indeterminate, based on the presence of this cluster of small dividing cells at their apex (Gehlot et al., 2012). Within the vast majority of these cells infection threads are present (Fig. 2b; arrow). Cells that contain infection threads can still divide (Fig. 2c; arrow). Further, the most distal cells in this cluster (Fig. 2b; arrow) contain infection threads. Both properties distinguish this cluster of cells from the meristem in indeterminate nodules of the IRLC group, for example in *Medicago* nodules. In addition, *I. argentea* nodules have some properties that have only been reported for determinate nodules. These are lenticels at their surface (Pankhurst and Sprent, 1975; Gehlot et al., 2012) (Fig. S1a), the absence of vacuoles in infected cells (Fig. 2a, b), a layer of uninfected cells between

the central tissue and nodule parenchyma, which is named boundary layer (Roussis et al., 1997) and the width of the infection threads is rather thin in comparison with indeterminate nodules (e.g. *Medicago*) (Fig. 2b, c; arrow). Taken together this questions whether *I. argentea* nodules are genuine indeterminate nodules.

***I. argentea* early nodule development is similar to that of determinate nodules**

The key characteristic of indeterminate nodules, as known from members of the IRLC clade, is the presence of a meristem at their apex that is formed at the primordium stage and is maintained during development. This meristem is formed from a cortical layer that is located closer to the root surface than the cortical cells that become infected in the primordium and will form several layers of infected cells in the proximal region of the nodule (Xiao et al., 2014). To study nodule development of *I. argentea*, root segments were collected at 10 and 11 days after inoculation (see Material and Methods). At these time points primordia/nodules at different stages of development are present and longitudinal sections were analysed. The youngest primordia show that anticlinal cell divisions were induced in the outermost cortical cell layer (Fig. 3a). The infected root hair associated with these primordia often has multiple infection threads. In slightly further developed primordia, several anticlinal and periclinal divisions have occurred in the outermost cortical cells and a small primordium has been formed (Fig. 3b). The infection threads have penetrated the primordium cells derived from the outer cortex (Fig. 3b). Anticlinal cell divisions have been induced in the pericycle and inner cortex and from these a vasculature will develop that connects the root and the primordium (Fig. 3c). Then the primordia developed into a young nodule, at which stage most of the cells in the central tissue were penetrated by infection threads and bacteria were released into cytoplasm of a few cells (Fig. 3c, arrowhead; Fig. 3e). Dividing infected cells at various stages of mitosis with infection threads were present (Fig. 3c, arrow; Fig. 3d). After release of bacteria, host cells stopped dividing (Fig. 3e). We tested whether mitotic activity correlated with the presence of relatively high auxin levels by using the auxin reporter DR5. Transgenic *I. argentea* nodules showed that DR5 was expressed in vascular bundles as well as in dividing cells with infection threads in the central tissue (Fig. 4a), and expression was absent after bacteria were released and cells no longer divided (Fig. 4b, c, d).

To determine whether *I. argentea* can form determinate nodules, 33 nodules were harvested 3 and 6 weeks after inoculation and nodules were completely sectioned. These nodules are heterogeneous with respect to their stage of development as in some cases primordia are still present. 6 nodules, of which all sections were analysed, had no meristem (or cluster of dividing cells) (Fig. 5b), which is similar to the cytology of determinate nodules.

The absence of a persistent meristem in determinate nodules results in peripheral tissues fully surrounding the central tissue. Therefore, we tested whether this is the case in *I. argentea* nodules. To visualise fully differentiated nodule parenchyma, we determined whether *I.*

argentea has an *ENOD2* gene. A putative *I. argentea* *ENOD2* gene (*IaENOD2*) was identified that is ~23 times upregulated in nodules in comparison to roots (Fig. S2c). The transcript contains a 930bp long open reading frame (ORF), encoding 309 amino acids (Fig. S2a, b) that is 50% identical to both *GmENOD2A* and *GmENOD2B* (Franssen et al., 1990;

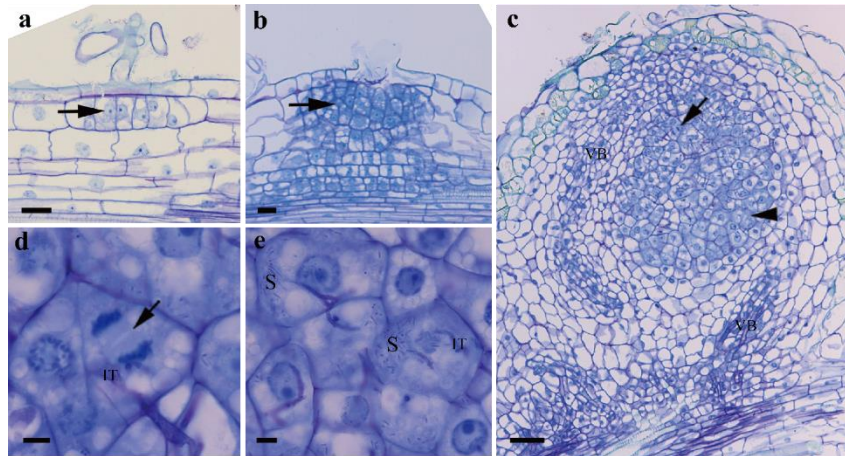


Fig. 3 Longitudinal section of *I. argentea* primordia and nodules formed at subsequent stages of development. (a) Anticlinal cell divisions are induced in the outer most cortical cells (arrow); (b) Anticlinal and periclinal cell divisions (arrow) have occurred in the outer most cortical cells and these form the nodule primordium; (c) A young nodule contains cells penetrated by infection threads (arrow) or with released bacteria in the cytoplasm (arrowhead); (d) Zoom in of the dividing infected cells with infection threads in (c) (arrow); (e) Zoom in of cells with released bacteria in the cytoplasm; VB, vascular bundle; IT: infection threads; S: symbiosome; Scale bars: a-b, 25 µm; c, 50 µm; d-e, 5 µm.

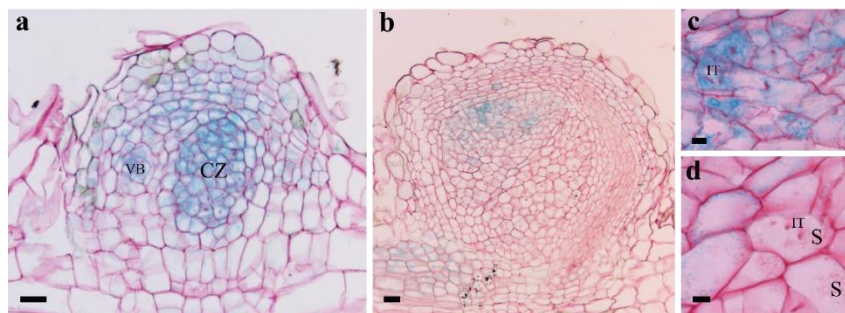


Fig. 4 DR5 promoter activity in *I. argentea* nodules. (a) GUS activity in central zone cells with infection thread and vascular bundles in young nodule; (b) GUS activity at the periphery of central zone in relatively old nodule; (c) Zoom in of relatively old nodule shows GUS activity only in cells with infection threads; (d) GUS activity is absent after bacteria are released. VB, vascular bundle; CZ: central zone; IT: infection thread; S: symbiosome; Scale bars: a-b, 25 µm; c-d, 5 µm.

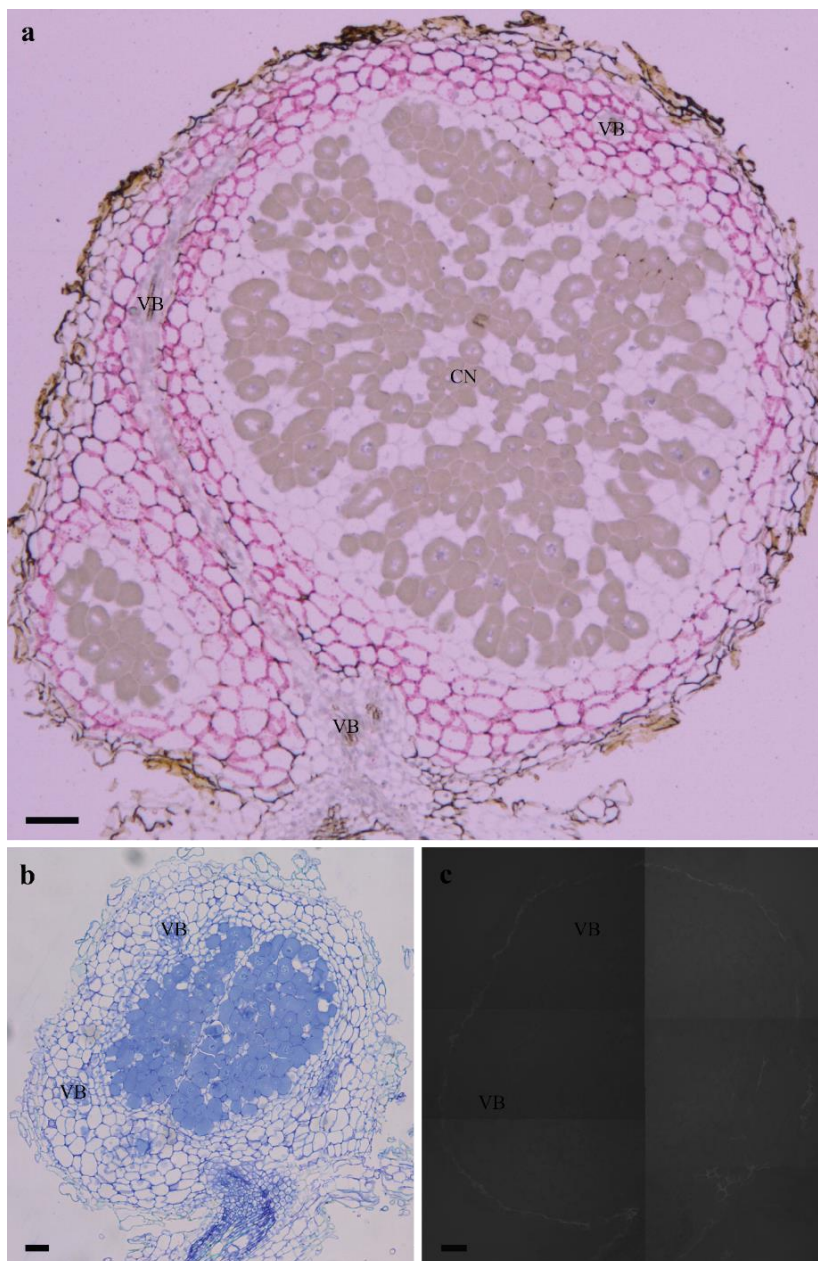


Fig. 5 Fully differentiated parenchyma and endodermis tissues in *I. argentea* nodules. (a) *IaENOD2* transcript is expressed in the nodule parenchyma that completely surrounds the central tissue in the nodule; (b) *I. argentea* nodule with no meristem at the apex; (c) Mature endodermis cells with Casparian strips completely surround the central tissue of the nodule. VB, vascular bundle; CZ: central zone; Scale bars: 75 μ m.

Schmutz et al., 2010), and 44% identical to *MtENOD2* from the more distantly related Medicago (Dickstein et al., 1988; Young et al., 2011). Phylogenetic analysis showed *IaENOD2* protein grouped together with GmENOD2A and GmENOD2B, SrENOD2 and MtENOD2 proteins are in another group (Fig. S2d), indicating that *I. argentea* is evolutionary more close to the determinate nodule forming soybean.

In situ hybridization shows that *IaENOD2* is specifically expressed in the nodule parenchyma (Fig. 5a, Fig. 7), consistent with being orthologous to *ENOD2* in other legume species. *In situ* hybridization in young nodules shows that *IaENOD2* is expressed in the nodule parenchyma that completely surrounds the central tissue (Fig. 5a). Mature nodule endodermis cells show auto-fluorescence under UV light due to the presence of Casparian strips. Visualisation of Casparian strips shows that in the nodules lacking a cluster of dividing cells, the mature endodermis also completely surrounds the central tissue (Fig. 5b, c). Further vascular bundles are present at the most distal part of the nodule (Fig. 5a-c), which is also a characteristic of determinate nodules (Sprent, 1980).

Thus, the early stages of *I. argentea* nodule development are very similar to those of determinate nodules and not of those of indeterminate nodules. This also raises the question how the cluster of dividing cells that is present in the older nodules is formed (Fig. 2a, b).

Identification of cells that form a cluster of dividing cells in mature *I. argentea* nodules

In longitudinal sections of the 33 3- and 6-week nodules, no dividing cells were observed in 6 of them (see above), whereas in 27 nodules, a cluster of dividing cells with infection threads was identified at their apex. The size of these clusters varied markedly. The smallest clusters were a single cell that just had divided into two daughter cells, and both cells could contain an infection thread (Fig. 6a, arrow). In bigger clusters, bacteria were released into the cytoplasm in some cells (Fig. S1b). Cells containing infection threads can still divide. However, we never observed dividing cells containing released bacteria. Some clusters were separated from the central tissue by one or two layers of non-dividing highly vacuolated cells (Fig. 6a), while others were directly adjacent to the infected cells of the central tissue (Fig. S1c). This suggests that the clusters can be formed from the boundary layer and the nodule parenchyma.

It therefore appears that *I. argentea* nodules elongate by creating secondary clusters of dividing cells that subsequently differentiate into nodule central tissues. This implies that the peripheral tissues also have to grow. To determine whether during elongation fully differentiated peripheral tissues surround the central tissue, we visualized the nodule parenchyma and endodermis. *In situ* hybridization on 4-week old nodules shows that *IaENOD2* is not expressed at the apex where the cluster of dividing cells is present (Fig. 7). Also Casparian strips are absent in this region (Fig. 6b, c). These observations indicate that

fully differentiated nodule parenchyma and endodermis cells dedifferentiate at the site where a cluster of dividing cells is formed.

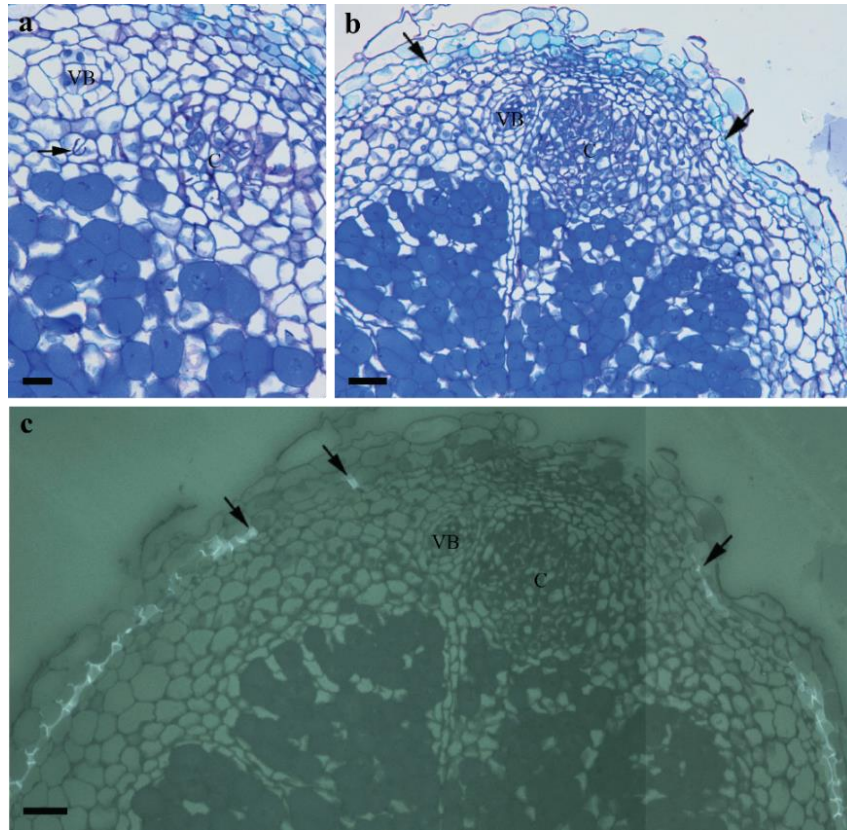


Fig. 6 Sections showing clusters of dividing cells at the apex of *I. argentea* nodules. (a) The smallest cluster of a single cell that just has divided into two daughter cells (arrow), infection thread is present in both cells; (b, c) Casparian strips (arrow) are absent at the apex of nodule where a cluster of dividing cells has been formed. C, cluster of dividing cells; VB, vascular bundle; Scale bars: 75 μ m.

Indigofera is one of the largest genera of flowering plants, with around 770 species, divided into four major clades on the basis of phylogenetic analyses (Schrire et al., 2009). *I. argentea* belongs to the Paleotropical clade. To determine whether the ability to form secondary clusters is common in *Indigofera* species, we investigated four additional species: *I. spinosa*, *I. linnaei*, *I. hochstetteri* from the Tethyan clade, and *I. tinctoria* from the Pantropical clade, respectively. In all these *Indigofera* species, the primordia were formed from the outermost cortical layer and these primordia did not form an apical meristem (Fig. S3-6). In all these species, a subset of nodules had no meristem at their apex nor clusters of dividing cells (sections of the complete nodule were analysed) (Fig. S3-6). In these nodules, the vascular

bundles were present in the most distal part of the nodule (Fig. S3-6). All these *Indigofera* species also form elongated nodules with clusters of dividing cells at their apex and also nodules with smaller clusters of dividing cells were found (Fig. S3-6). In *I. tinctoria* from the Pantropical clade, infection threads are present in the most distal cells in cluster of dividing cells (Fig. S3), which is similar to *I. argentea* nodules. In the three *Indigofera* species from the Tethyan clade, infection threads only exist for a very short period. In the clusters of dividing cells, released bacteria were detectable and these cells were still mitotically active (Fig. S4-6). This is similar to *Lupinus* nodules in which it has previously been reported that cells remain mitotically active after release of rhizobia into the cytoplasm (Gonzalez-Sama et al., 2004; Fedorova et al., 2007).

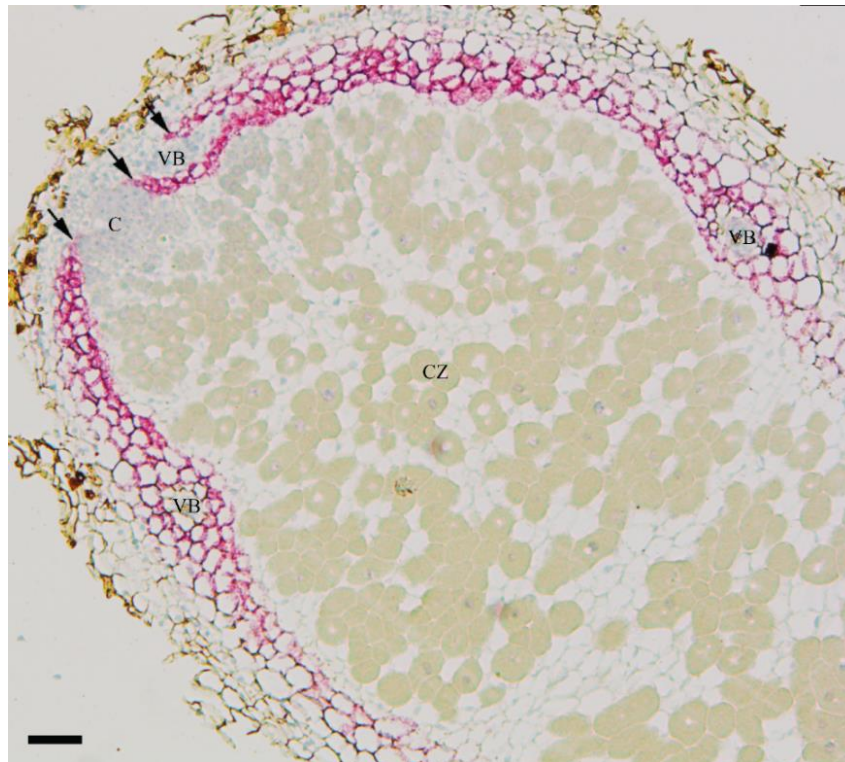


Fig. 7 Localization of *IaENOD2* transcripts by *in situ* hybridization. *IaENOD2* transcripts are absent in the tissue adjacent to the cluster of dividing cells. C, cluster of dividing cells; VB, vascular bundle; CZ: central zone; Scale bar: 75 μ m.

Thus, all studied *Indigofera* species form determinate nodules and have the ability to form secondary clusters of dividing cells. This implies that within the millettiods clade *Tephrosia*

remains the only known genus that has been reported to form indeterminate nodules (Fig. 1). This prompted us to re-study *Tephrosia*.

***Tephrosia* forms determinate nodules**

Three *Tephrosia* species (*T. villosa*, *T. wallichii*, *T. purpurea*) were studied. In these species nodule primordia were induced in the outermost cortical layer (Fig. 8a; Fig. S7-8), and an apical meristem was never observed in these primordia. Further, a subset of the nodules lacks dividing cells at their apex (all sections of complete nodules analysed) (Fig. 8b; Fig. S7-8). In these nodules, vascular bundles are present in the most distal region of the nodule. Infection threads in these three *Tephrosia* species only exist for a very short period. Clusters of small dividing cells with released bacteria were detectable (Fig. 8c, d; Fig. S7-8), even though they are not located in boundary/parenchyma tissue like in *I. argentea*, *I. tinctoria*, or soybean nodules, but more similar to nodules of the three *Indigofera* species from the Tethyan clade (Fig. 8d; Fig. S4-6).

Based on these observations we conclude that *Tephrosia* species form determinate nodules in a similar way as *Indigofera* species.

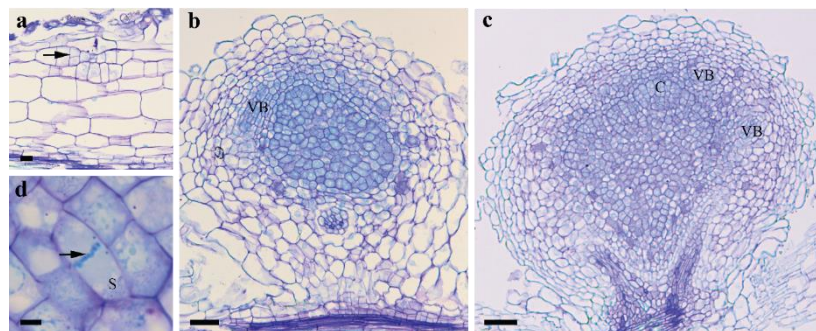


Fig. 8 Longitudinal sections of primordia and nodules formed on *T. villosa*. (a) Primordia were induced in the outermost cortical cell layers (arrow); (b) Determinate nodules with no meristem at the apex formed on *T. villosa*; (c) Cluster of dividing cells at the apex of nodule; (d) Zoom in of dividing cells (arrow) with intracellular bacteria in nodule showed in (c). VB, vascular bundle; C, cluster of dividing cells; S: symbiosome; Scale bars: a, 25 µm; b, 50 µm; c, 75 µm; d, 5 µm.

Discussion

We showed that *Indigofera* species make nodules that do not have a persistent meristem and they are formed from the outermost root cortical cell layer (Fig. 3; Fig. S3-6). Therefore, we classify them as determinate nodules. They were previously considered to be indeterminate, because mature nodules can have a cluster of dividing cells at their apex (Oono et al., 2010; Gehlot et al., 2012). However, this cluster of dividing cells is the result of mitotic reactivation

of nodule parenchyma and boundary cells by which a secondary nodule meristem is formed (Fig. 2a, b; Fig. 6a, b; Fig. S3-6). Also, *Tephrosia* nodules, that previously also were classified as indeterminate (Oono, et al., 2010), were shown to be determinate (Fig. 8a, b). This implies that all analysed species belonging to the tribe Indigofereae and to the millettoids clade form determinate nodules. We hypothesise that the common ancestor of IRLC, robinoids clade, *Indigofera*, and millettoids clade (Fig. 1) formed determinate nodules and the indeterminate nodule type evolved in an ancestor of the IRLC and robinoids clade.

The cluster of dividing cells that can occur at the apex of mature *I. argentea* nodules is formed by mitotic reactivation of fully differentiated nodule cells. We reach this conclusion based on the observation that most young nodules have no apical meristem. Further, the cluster of dividing cells varies in size, with the smallest cluster being formed by a single cell that has divided once (Fig. 6a). These clusters of dividing cells can easily be recognized as they contain numerous infection threads (Fig. 2b; Fig. 6a, b). These occur even in the most distal cells within these clusters and cells containing infection threads are still able to divide (Fig. 2c; Fig. 3d). The occurrence of infection threads as well as their ontogeny make these dividing cells markedly different from the meristem cells in indeterminate nodules (Xiao et al., 2014). Also, the 4 other *Indigofera* species, that we studied, form determinate nodules and can form secondary clusters of dividing cells.

Indigofera nodules lose their primary meristematic cells at a relatively early stage and this results in peripheral tissues that completely surround the central tissue. This was shown to be the case for endodermis and nodule parenchyma (Fig. 5a-c). Also during *Tephrosia* nodule development an apical meristem is not formed in primordia and mitotic activity is not maintained. Therefore, *Tephrosia* also forms determinate nodules.

At the site where a secondary cluster of dividing cells is formed the nodule parenchyma and endodermis are dedifferentiated, as *ENOD2* expression and Casparian strips, respectively, are lost (Fig. 6b, c; Fig. 7). This is probably required for growth into an elongated nodule (Fig. 2a; Fig. S1a). It seems probable that cells derived from dedifferentiated nodule parenchyma and boundary layer form the secondary cluster of dividing cells.

We showed that *Indigofera* and *Tephrosia* species form secondary clusters of dividing cells. This ability to form secondary clusters of dividing cells could be a trait that is more widespread among legumes that form determinate nodules. Soybean nodules are iconic determinate nodules. However, it has been shown that during recovery after salt stress, clusters of dividing cells are formed from fully differentiated nodule parenchyma cells. These cells become infected by rhizobia (James et al., 1993). It has been shown that in *Sesbania* nodules mitotically active cells can be present in peripheral tissues. This was demonstrated by *in situ* hybridization with a mitotic cyclin probe (Fernández-lópez et al., 1998). Whether such mitotically activated cells can become infected is not known. We hypothesise that under

specific conditions, formation of secondary clusters of dividing cells can be induced in determinate nodules. In some species of genera such as *Indigofera*, the “normal” growth conditions in the lab are sufficient to trigger this. However, in other species stress or other environmental factors, that in most cases are still unknown, are probably needed. The soybean experiments indicate that this can involve very specific conditions as in this species the formation of a secondary cluster of dividing cells, after transient salt stress, even depended on the light quality (James et al., 1993).

Determinate nodules of species of *Indigofera* and the millettoids clade (e.g. *Tephrosia*, soybean) are formed from the outermost cortical cell layer (Fig. 3a-c; Fig. 8a; Fig. S3-8). In contrast, indeterminate nodules of IRLC species like *Medicago*, are formed from inner cortical cell layers. Detailed characterization of early steps of *Medicago* nodule (primordium) development (Xiao et al., 2014) shows that the meristem is formed from the middle cortical cell layer (3th layer). This cell layer divides later than the inner cortical cell layers (4th and 5th layer). The infection threads enter cells derived from the inner cortical cells, but it can only pass the meristem-forming middle cortical layers before it is mitotically activated or when only a few anticlinal divisions have occurred. The cells of the meristem of these indeterminate nodules are not penetrated by an infection threads, whereas the daughter cells are. So the formation of a non-infected apical meristem seems to require a cortical cell layer that is located more outward than the cortex-derived cells that become infected (and form the proximal part of the nodule). This strongly suggests that in cases where cells derived from the outermost cortical cell layer become infected, a non-infected apical meristem cannot be formed in the nodule primordium. To our knowledge it also has never been shown that an apical meristem is formed in a nodule primordium that is formed from the outermost cortical cell layer. This idea underlines the conclusion that *Indigofera* and *Tephrosia* species form determinate nodules as in both genera nodule primordia are formed from cells derived from the outermost cortical cell layer.

The formation of a secondary cluster of dividing nodule cells is a fundamentally different developmental process than the formation of a non-infected persistent apical meristem that is formed at the primordium stage. The indeterminate growth that drives nodule elongation in genuine indeterminate nodules (e.g. *Medicago*) or determinate nodules (e.g. *Indigofera*) is the result of convergent evolution. To understand legume nodule evolution, it will be important to distinguish “indeterminate nodule growth” driven by a primary persistent meristem from that controlled by a secondary cluster of dividing cells. The latter occurs in determinate nodules. The wrong qualification of *Indigofera* and *Tephrosia* nodules indicates that this might have happened in (many) other legume species. This is almost certainly the case for species that form nodule primordia from outer cortical cells. This is for example the case in *Lupinus* where primordia originate from the outer cortex (Tang et al., 1992; James et al., 1997; Gonzalez-Sama et al., 2004; Lotocka, 2008), but nodules are classified as indeterminate. *Lupinus* has clusters of dividing cells and these cells remain mitotically active

after release of rhizobia (Gonzalez-Sama et al., 2004; Fedorova et al., 2007). This seems very similar to *Tephrosia* (Fig. 8c). Therefore, to understand nodule evolution it will be important to reanalyse nodule development in other tribes.

Our studies show that all studied genera within the milletioids clade and Indigoferaeae form determinate nodules. Further, *Lupinus* nodule development might be very similar to *Tephrosia* and *Indigofera* nodule development. Therefore, we assume that the common ancestor of the 4 selected groups (node 15) formed determinate nodules. This suggests that the genuine indeterminate nodule type first evolved within the IRLC clade. The robinoids clade is rather puzzling as it contains genera (e.g. *Robinia*) that are reported to make indeterminate nodules as well as genera (e.g. *Lotus*) that form determinate nodules. It is likely that the development of determinate nodules (*Lotus*, *Sesbania*) within this tribe differs from that of the other determinate nodules within the 4 selected groups. The determinate nodules of *Sesbania* as well *Lotus* are formed from cells derived from the middle cortical cell layers, whereas all other determinate nodules are formed from the outermost cortical cell layers. Further, *Sesbania* can even develop indeterminate as well as determinate nodules (Fernández-lópez et al., 1998; Goormachtig et al., 1998). Taken together it seems equally well possible that within the robinoid clade an evolution from indeterminate back to determinate nodule type occurred or that indeterminate nodules evolved independently.

Materials and Methods

Plant material, rhizobial strain isolation, and root inoculation

This study included 5 *Indigofera* species and 3 *Tephrosia* species (Table 1), of which *I. argentea* and *I. spinosa* were collected from Jizan desert, Saudi Arabia. *I. linnaei*, *I. hochstetteri*, *I. tinctoria*, *T. wallichii*, *T. villosa*, and *T. purpurea* were provided by Hukam S. Gehlot from Jai Narain Vyas University, Jodhpur, India (Gehlot et al., 2012). All *Indigofera* species and *Tephrosia* species were grown in sand from Jizan desert. Rhizobia were isolated as described in Chapter 2 and were used to inoculate plants (Table 1). Seeds germination and root inoculation were performed as described in Chapter 2.

Tissue embedding, sectioning and section staining

Root segments were stained with 0.01% methylene blue (w/v) to detect primordia with a binocular. Primordia and nodules were fixed with 5% glutaraldehyde (v/v) in 0.05M sodium phosphate buffer (pH 7.0) for 15 hours at 4°C, after vacuum infiltration for 30 min. Subsequently, the material was dehydrated in an ethanol series and then embedded in Technovit 7100 (Heraeus Kulzer) according to the manufacture's protocol. Longitudinal sections (7 µm thick) were made by using a RJ2035 microtome (Leica Microsystems, Rijswijk, the Netherlands) and stained for 3 min in 0.05% Toluidine Blue O (w/v). Sections

were examined with a DM5500B microscope equipped with a DFC425C camera (Leica Microsystems, Wetzlar, Germany).

Table 1 *Indigofera* and *Tephrosia* species, as well as rhizobia isolates used for inoculation

Species	Rhizobial strains
<i>Indigofera</i>	
<i>I. argentea</i>	<i>Bradyrhizobium elkani</i> , SA281
<i>I. spinosa</i>	<i>Bradyrhizobium elkani</i> , IS003
<i>I. hochstetteri</i>	<i>Bradyrhizobium elkani</i> , IH010
<i>I. linnaei</i>	<i>Bradyrhizobium elkani</i> , IL001
<i>I. tinctoria</i>	<i>Ensifer kostiense</i> , IT001
<i>Tephrosia</i>	
<i>T. wallichii</i>	<i>Ensifer saheli</i> , TW010
<i>T. villosa</i>	<i>Ensifer sp.</i> , TV001
<i>T. purpurea</i>	<i>Bradyrhizobium elkanii</i> , TP005

RNA isolation and transcriptome analysis

RNA isolation and *de novo* transcriptome analysis was performed by as described (Chapter 2). A putative *Indigofera* ENOD2 orthologue (*IaENOD2*, Fig. S2a) was identified using *GmENOD2A* (X16875.1), *GmENOD2B* (X16876.1) and *MtENOD2* (XM_013613246) as query sequences.

RNA *in situ* hybridization

Nodules were fixed with 5% glutaraldehyde in 0.05M sodium phosphate buffer (pH 7.0) (RNase free) and embedded in paraffin (Paraplast X-tra, McCormick Scientific). Nodule sections of 5 µm thick were prepared with a RJ2035 microtome. *IaENOD2* probe set containing 20 oligonucleotide pairs was designed and produced by Affymetrix based on the *IaENOD2* sequence and covered the region 30-905 nt. *In situ* hybridization was conducted according to the user manual for ViewRNA ISH Tissue 2-plex Assay (<http://www.panomics.com/UserDocs>). Slides were analysed with an AU5500B microscope equipped with a DFC425c camera (both Leica).

Expression analysis and histochemical GUS staining

Expression analysis of nodules containing promoter-GUS fusions were performed as described in Chapter 2.

References

- Bergersen, F.J., and Briggs, M.J. (1958). Studies on the bacterial component of soybean root nodules: Cytology and organization in the host tissue. *J. gen. Microbiol.* 19: 482-490.
- Brundrett, M.C., Enstone, D.E., and Peterson, C.A. (1988). A berberine-aniline blue fluorescent staining procedure for suberin, lignin, and callose in plant-tissue. *Protoplasma* 146(2-3): 133-142.
- Dickstein, R., Bisseling, T., Reinhold, V.N., and Ausubel, F.M. (1988). Expression of nodule-specific genes in alfalfa root-nodules blocked at an early stage of development. *Genes & Development* 2(6): 677-687.
- Doyle, J.J. (2011). Phylogenetic perspectives on the origins of nodulation. *Molecular Plant-Microbe Interactions* 24(11): 1289-1295.
- Fedorova, E.E., De Felipe, M.R., Pueyo, J.J., and Lucas, M.M. (2007). Conformation of cytoskeletal elements during the division of infected *Lupinus albus* L. nodule cells. *J Exp Bot* 58(8): 2225-2236.
- Fernández-lópez, M., Goormachtig, S., Gao, M., D'Haeze, W., Van Montagu, M., and Holsters, M. (1998). Ethylene-mediated phenotypic plasticity in root nodule development on *Sesbania rostrata*. *Proc Natl Acad Sci U S A* 95(21): 12724-12728.
- Franssen, H.J., Thompson, D.V., Idler, K., Kormelink, R., Van Kammen, A., and Bisseling, T. (1990). Nucleotide sequence of two soybean *ENOD2* early nodulin genes encoding Ngm-75. *Plant Mol Biol* 14(1): 103-106.
- Gehlot, H. S., Panwar, D., Tak, N., Tak, A., Sankhla, I.S., Poonar, N., Parihar, R., Shekhawat, N.S., Kumar, M., Tiwari, R., Ardley, J., James, E.K., and Sprent, J.I. (2012). Nodulation of legumes from the Thar desert of India and molecular characterization of their rhizobia. *Plant and Soil* 357(1-2): 227-243.
- Gonzalez-Sama, A., Lucas, M.M., De Felipe, M.R., and Pueyo, J.J. (2004). An unusual infection mechanism and nodule morphogenesis in white lupin (*Lupinus albus*). *New Phytologist* 163(2): 371-380.
- Goormachtig, S., Van Montagu, M., and Holsters, M. (1998). The early nodulin gene *ENOD2* shows different expression patterns during *Sesbania rostrata* stem nodule development. *Molecular Plant-Microbe Interactions* 11(3): 237-241.
- Hirsch, A.M. (1992). Developmental biology of legume nodulation. *New Phytologist* 122(2): 211-237.
- James, E.K., Sprent, J.I., Hay, G.T., and Minchin, F.R. (1993). The effect of irradiance on the recovery of soybean nodules from sodium chloride-induced senescence. *Journal of Experimental Botany* 44(263): 997-1005.
- James, E.K., Minchin, F.R., Iannetta, P.P.M., and Sprent J.I. (1997). Temporal relationships between nitrogenase and intercellular glycoprotein in developing white lupin nodules. *Annals of Botany* 79: 493-503.
- Lauridsen, P., Franssen, H., Stougaard, J., Bisseling, T., and Marcker, K.A. (1993). Conserved regulation of the soybean early nodulin *ENOD2* gene promoter in determinate and indeterminate transgenic root nodules. *Plant J* 3(3): 483-492.
- Lavin, M., Pennington, R.T., Klitgaard, B.B., Sprent, J.I., De Lima, H.C., and Gasson, P.E. (2001). The dalbergioid legumes (Fabaceae): Delimitation of a pantropical monophyletic clade. *American Journal of Botany* 88(3): 503-533.
- Lemaire, B., Dlodlo, O., Chiphango, S., Stirton, C., Schrire, B., Boatwright, J.S., Honnay, O., Smets, E., Sprent, J., James, E.K., and Muasya, A.M. (2015). Symbiotic diversity, specificity and distribution of rhizobia in native legumes of the Core Cape Subregion (South Africa). *FEMS Microbiology Ecology* 91(2): 1-17.
- Lotocka, B. (2008). Vascular system within developing root nodules of *Lupinus luteus* L. part 1. juvenile stage. *Acta Biologica Cracoviensia Series Botanica* 50(1): 79-88.
- Madsen, L.H., Tirichine, L., Jurkiewicz, A., Sullivan, J.T., Heckmann, A.B., Bek, A.S., Ronson, C.W., James, E.K., and Stougaard, J. (2010). The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat Commun* 1:10.
- Ndoye, I., De Billy, F., Vasse, J., Dreyfus, B., and Truchet, G. (1994). Root nodulation of *Sesbania rostrata*. *J Bacteriol* 176(4): 1060-1068.
- Oono, R., Schmitt, I., Sprent, J.I., and Denison, R.F. (2010). Multiple evolutionary origins of legume traits leading to extreme rhizobial differentiation. *New Phytologist* 187(2): 508-520.
- Pankhurst, C.E., and Sprent, J.I. (1975). Surface features of soybean root nodules. *Protoplasma* 85(1): 85-98.
- Remigi, P., Zhu, J., Young, J.P.W., and Masson-Boivin, C. (2016). Symbiosis within symbiosis: Evolving nitrogen-fixing legume symbionts. *Trends in Microbiology* 24(1): 63-75.
- Roussis, A., Papadopoulou, K., and Katinakis, P. (1997). NOD3, a novel late nodulin gene from soybean is expressed in the infected cells and the nodule parenchyma. *Journal of Experimental Botany* 48(310): 1011-1017.
- Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463(7278): 178-183.

- Schrire, B.D., Lavin, M., Barker, N.P., and Forest, F. (2009).** Phylogeny of the tribe Indigofereae (Leguminosae-Papilionoideae): Geographically structured more in succulent-rich and temperate settings than in grass-rich environments. *Am J Bot* 96(4): 816-852.
- Sprent, J.I. (1980).** Root nodule anatomy, type of export product and evolutionary origin in some Leguminosae. *Plant Cell and Environment* 3(1): 35-43.
- Sprent, J.I. (2001).** Formation, structure and function of nodules. *Nodulation in legumes*. Kew, Royal Botanic Gardens: 35.
- Tang, C.X., Robson, A.D., Dilworth, M.J., and Kuo, J. (1992).** Microscopic evidence on how iron deficiency limits nodule initiation in *Lupinus angustifolius* L. *New Phytologist* 121(3): 457-467.
- Turgeon, B.G., and Bauer, W.D. (1982).** Early events in the infection of soybean by *Rhizobium japonicum*. Time course and cytology of the initial infection process. *Canadian Journal of Botany* 60(2): 152-161.
- Van de Wiel, C., Scheres, B., Franssen, H., Vanlierop, M.J., Van Lammeren, A., Van Kammen, A., and Bisseling, T. (1990a).** The early nodulin transcript *ENOD2* is located in the nodule parenchyma (inner cortex) of pea and soybean root-nodules. *Embo Journal* 9(1): 1-7.
- Van de Wiel, C., Norris, J.H., Bochenek, B., Dickstein, R., Bisseling, T., and Hirsch, A.M. (1990b).** Nodulin gene-expression and *ENOD2* localization in effective, nitrogen-fixing and ineffective, bacteria-free nodules of alfalfa. *Plant Cell* 2(10): 1009-1017.
- Vijn, I., Martinezabarca, F., Yang, W.C., Dasneves, L., Vanbrussel, A., Van Kammen, A., and Bisseling, T. (1995).** Early nodulin gene expression during Nod Factor-induced processes in *Vicia sativa*. *Plant Journal* 8(1): 111-119.
- 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(18): 3517-3528.
- Yang, G., Zhou, R., Tang, T., and Shi, S. (2008).** Simple and efficient isolation of high-quality total RNA from *Hibiscus tiliaceus*, a mangrove associate and its relatives. *Prep Biochem Biotechnol* 38(3): 257-264.
- Young, N.D., Debelle, F., Oldroyd, G.E.D., Geurts, R., Cannon, S.B., Udvardi, M.K., Benedito, V.A., Mayer, K.F.X., Gouzy, J., Schoof, H., et al. (2011).** The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480(7378): 520-524.

Supplementary Figures

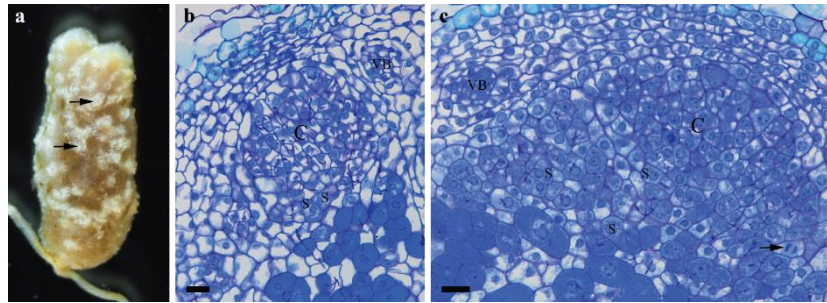


Fig. S1 *I. argentea* nodules. (a) Elongated nodule with lenticels (arrow) in *I. argentea*; (b) Bacteria are released in some cells of bigger cluster of dividing cells; (c) Cluster of dividing cells just adjacent to the central tissue, dividing infected cell is present (arrow). VB, vascular bundle; C, cluster of dividing cells; S: symbiosome; Scale bars: b-c, 25 µm.

(a) 1 10 20 30 40 50
 | | | | |
 CGCCGGGATCACAAGCCAGATATGTCTTCTTCTATACACTCCTCAATAG
 CTTTGCTCTACTTGGAGTAATGGTGACTCAAGTCACCACTCCAGTCTTG
 GGGAACCTTTTACAACCCACCCACTTACGAGCCTCCACCCTCCTATGTCTA
 TGAGCCACCACCACCGGTGTATAAGCCACCATTCTACCCACCACCAATTT
 ACCACCCTCCACATGAGAAGCCACCTCCCGAGAAACCACGTACGAGAAA
 CCGCTCAGGAGAAGCCTCCGTATCAAAAACCACCTACGAGAAGCCTCC
 GTATGAGAAGCCGCTCAGGAGAAACCACCTCATGAAAAGCCTCCATACG
 AGAAGCCACCTCATGAGAAGCCCATATGAGAAGCCACCTCATGAAAA
 CCCCCATACGAGAAACCACCACATGAGAAGCCACCACATGAGAAGCCACC
 TTATGAGAAGCCACCATATGAAAAGTCACCCCGAGAATACCAACCACCTC
 CAGAAATACAGCCTCCTCAGGAGAAGCCACCACCGGAATACGAGCCACCA
 CCAAAGTATGAGCCACCACATGAAAAGCCCCGCGCAGAGTATGAGCCTCC
 TCACGAGAAGCCACCCACAGGTACGAACCCACCACCATATGAACCAC
 CTTATGAGAAGCCACCACAGTGTATGAACCGCCATATGAGAAGCCACCA
 CCAGTGTACGAACCGCCATATGAGAAGCCACCACAGTGTACGAACCGCC
 ATATGAGAAGCCACCACAGTGTACGAACCGCCATATGAGAAGCCACCAC
 CAGTGTACGAACCGCCATATGAGAAGCCACCACCAATGTACGAACCGCCT
 TATGAGAAGCCACCACAGTTTATGAACCACCACCTCATGAGAAGCCACC
 CTTTTACAACCTCCTTATCAGAAGCCACCATCCAAAGGCAAAAAGAAGT
 AGAAACCAGTCATGCATGTGACATATTTAATTTCTTTACTAGTCTAAA
 GTAATAAAGCAGTCCTTATCCTATAAATTAAAGTCTGTGTTTGTCTG
 TTCAGGAAGCAAGAAGCAGTTTTTACTATATGTCAATTACAATGTTGTCT
 TTTGGGCTCTGTCATTGTTCTAAAAATAAAGGCTCTCTATACGCCTGTG
 TGTAAATATCTTTCCATTTCGTTTCAGCAACCCCATAAATGTAATAATAT
 TGTTCCCTTTTCTATGGCTAATTAAATACGATACTACTTTTTTC

(b) 1 10 20 30 40 50
 | | | | |
 MSSSIHSSIALLLLGVMTQVTTTPVLGNFYNPPTYEPPPSYVYEPPIPVY
 KPPFYPPPIYHPHEKPPPEKPPYEKPPHEKPPYQKPPYEKPPYEKPPHE
 KPPHEKPPYEKPPHEKPPYEKPPHEKPPYEKPPHEKPPHEKPPYEKPPYE
 KSPPEYQPPPEYQPPHEKPPPEYEPPIKYEPHEKPPPEYEPHEKPPPE
 YEPPIYEPPIYEPPIYEPPIYEPPIYEPPIYEPPIYEPPIYEPPIYEPPIY
 YEPPIYEPPIYEPPIYEPPIYEPPIYEPPIYEPPIYEPPIYEPPIYEPPIY
 KPPSKGKKK

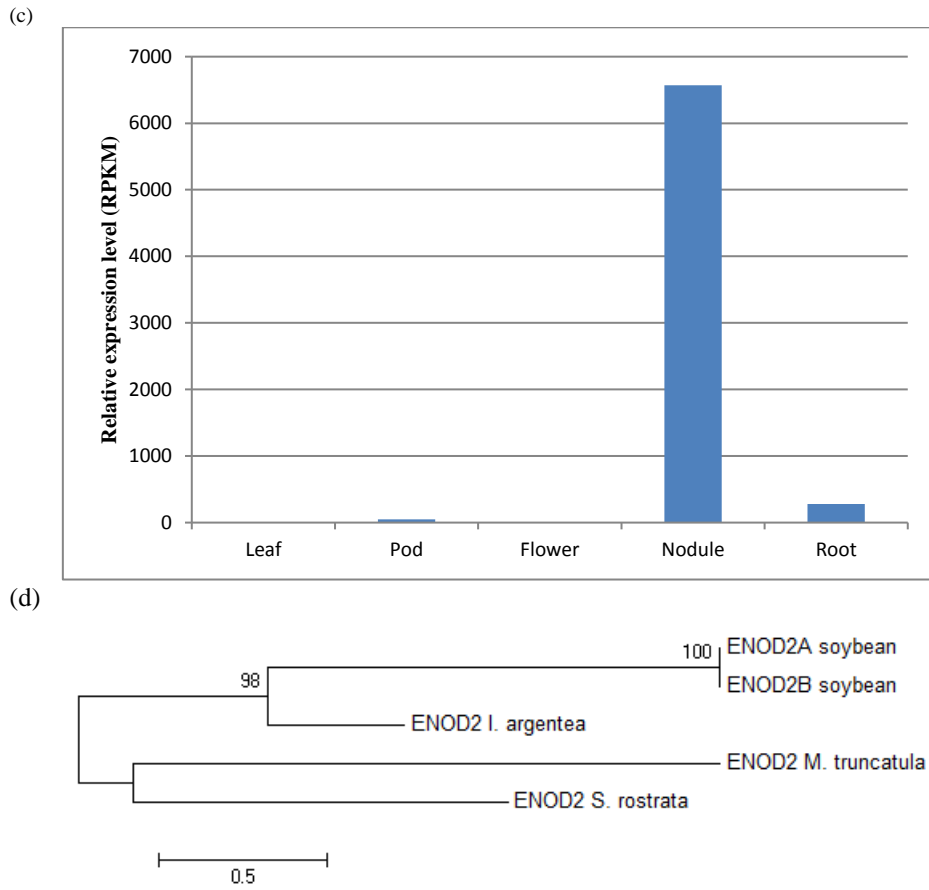


Fig. S2 (a) Nucleotide sequence of *IaENOD2* transcript, a length of 930bp open reading frame (ORF) from site 23 to 952 was underlined; (b) Translated amino acid sequence of *IaENOD2* ORF, including eleven PPHEK repeats and one PPPHEK, fifteen PPYEK repeats, six PPPVYE repeats and six variants with one amino acid replacement (E/K/T/M); (c) *IaENOD2* transcript expression level in nodule is about 23-fold comparing to root (RPKM); (d) Phylogenetic tree of *ENOD2* orthologues from four species indicating *IaENOD2* is more related to two soybean *ENOD2*.

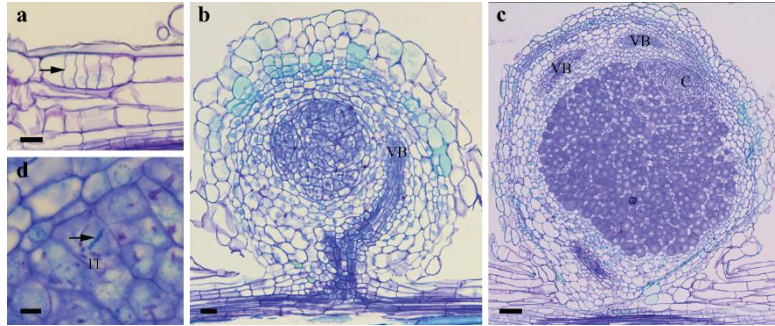


Fig. S3 Longitudinal sections of primordia and nodules formed on *I. tinctoria*. (a) Primordia were induced in the outer most cortical cell layers (arrow); (b) Determinate nodules with no meristem at the apex formed on *T. villosa*; (c) Cluster of dividing cells at the apex of nodule; (d) Zoom in of dividing cells (arrow) with intracellular bacteria in nodule showed in (c). VB, vascular bundle; C, cluster of dividing cells; IT: infection thread; Scale bars: a-b, 25 μ m; c, 75 μ m; d, 7.5 μ m.

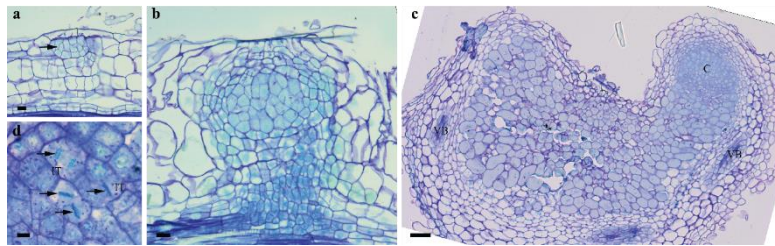


Fig. S4 Longitudinal sections of primordia and nodules formed on *I. spinosa*. (a) Primordia were induced in the outer most cortical cell layers (arrow); (b) Determinate nodules with no meristem at the apex formed on *T. villosa*; (c) Cluster of dividing cells at the apex of nodule; (d) Zoom in of dividing cells (arrow) with intracellular bacteria in nodule showed in (c). VB, vascular bundle; C, cluster of dividing cells; IT: infection thread; Scale bars: a-b, 25 μ m; c, 75 μ m; d, 7.5 μ m.

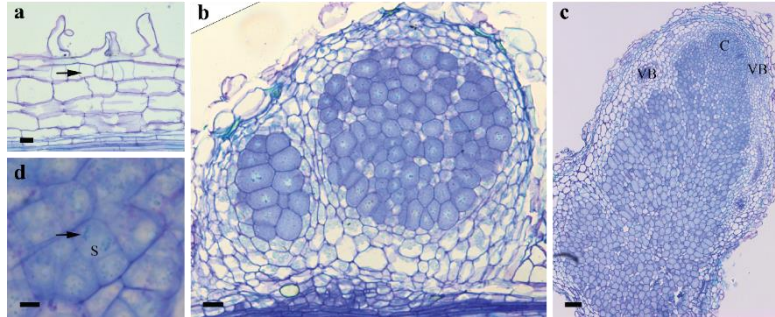


Fig. S5 Longitudinal sections of primordia and nodules formed on *I. linmaei*. (a) Primordia were induced in the outer most cortical cell layers (arrow); (b) Determinate nodules with no meristem at the apex formed on *T. villosa*; (c) Cluster of dividing cells at the apex of nodule; (d) Zoom in of dividing cells (arrow) with intracellular bacteria in nodule showed in (c). VB, vascular bundle; C, cluster of dividing cells; S: symbiosome; Scale bars: a-b, 25 µm; c, 75 µm; d, 7.5 µm.

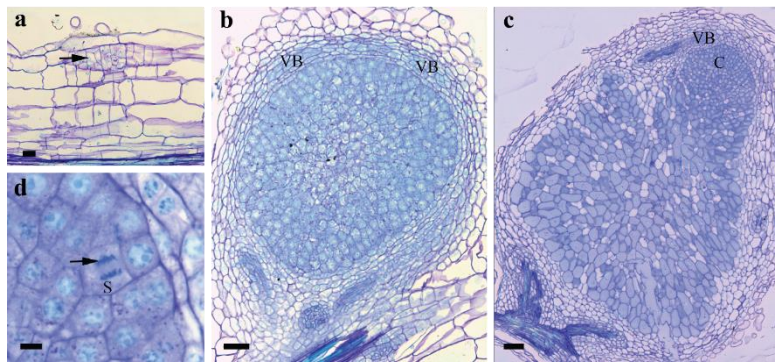


Fig. S6 Longitudinal sections of primordia and nodules formed on *I. hochstetteri*. (a) Primordia were induced in the outer most cortical cell layers (arrow); (b) Determinate nodules with no meristem at the apex formed on *T. villosa*; (c) Cluster of dividing cells at the apex of nodule; (d) Zoom in of dividing cells (arrow) with intracellular bacteria in nodule showed in (c). VB, vascular bundle; C, cluster of dividing cells; S: symbiosome; Scale bars: a-b, 25 µm; c, 75 µm; d, 7.5 µm.

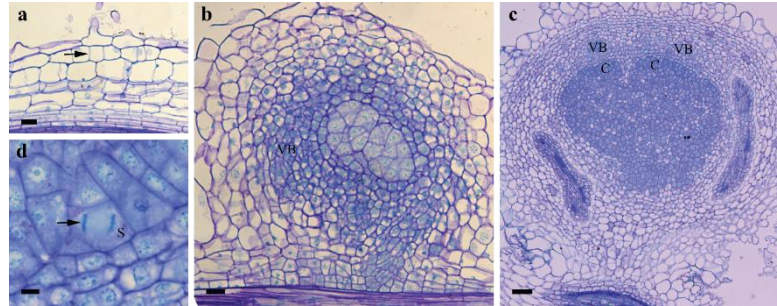


Fig. S7 Longitudinal sections of primordia and nodules formed on *T. wallichii*. (a) Primordia were induced in the outer most cortical cell layers (arrow); (b) Determinate nodules with no meristem at the apex formed on *T. villosa*; (c) Cluster of dividing cells at the apex of nodule; (d) Zoom in of dividing cells (arrow) with intracellular bacteria in nodule showed in (c). VB, vascular bundle; C, cluster of dividing cells; S: symbiosome; Scale bars: a-b, 25 μ m; c, 75 μ m; d, 7.5 μ m.

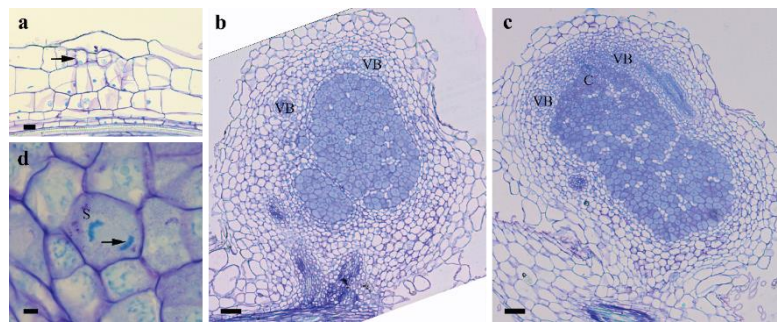


Fig. S8 Longitudinal sections of primordia and nodules formed on *T. purpurea*. (a) Primordia were induced in the outer most cortical cell layers (arrow); (b) Determinate nodules with no meristem at the apex formed on *T. villosa*; (c) Cluster of dividing cells at the apex of nodule; (d) Zoom in of dividing cells (arrow) with intracellular bacteria in nodule showed in (c). VB, vascular bundle; C, cluster of dividing cells; S: symbiosome; Scale bars: a, 25 μ m; b-c, 75 μ m; d, 5 μ m.

Chapter 5

The role of NCRs in bacteroid differentiation in *Indigofera argentea* root nodules

Guiling Ren¹, Arjan van Zeijl¹, Robin van Velzen¹, Peter Mergaert², Henk Franssen¹, Carolien Franken¹, Jan Verver¹, Ton Bisseling¹

¹ Laboratory of Molecular Biology, Plant Sciences Group, Wageningen University and Research Centre, Wageningen, The Netherlands

² Institute for Integrative Biology of the Cell, Centre National de la Recherche Scientifique, UMR 9198, 91198, Gif-sur-Yvette, France.

Abstract

Rhizobia induce nodules on legume roots in which rhizobia differentiate into an endosymbiotic form named bacteroid. In legume species of the IRLC clade (e.g. pea, *Medicago*), bacteroids undergo a physiological and morphological transformation, which is induced by Nodule-specific Cysteine Rich peptides (NCRs). This involves terminal bacteroid differentiation and an increase in bacteroid size and ploidy. The role of NCRs in inducing bacteroid differentiation was considered to be unique for IRLC legumes. However recently it was shown that in legumes of the dalbergioids clade (e.g. *Aeschynomene*), bacteroid undergo a similar transformation, which is also controlled by NCR-like peptides. Here, we show that bacteroids in *Indigofera argentea* nodules, from the indigoferoids clade, are also elongated or branched and polyploid, similar to bacteroids in nodules of IRLC and dalbergioid legumes. Analysis of *I. argentea* transcriptomes, identified a group of NCR peptides. A possible role of NCRs in bacteroid differentiation in *I. argentea* nodules and the independent evolution of NCRs in the IRLC, dalbergioids and indigoferoids clades is discussed.

Introduction

Rhizobia can establish a N_2 -fixing nodule symbiosis with legumes. This symbiosis occurs in all three legume sub-families, but is especially prominent in the Papilionoideae subfamily. This subfamily contains the model legumes *Lotus* and *Medicago*, as well as most legume crops like pea, cowpea, clover and soybean. In legume nodules, the rhizobia are hosted intracellularly in specialized infected nodule cells, where they differentiate into N_2 -fixing bacteroids. In species of the IRLC clade (Fig. 1), e.g. *Medicago* and pea, bacteroids markedly enlarge, which is correlated with endoreduplication of the bacteroids and a loss of the ability to return to the free-living bacterial state. This bacteroid enlargement is triggered by Nodule-specific Cysteine Rich (NCR) peptides in species of the IRLC clade. However, more recently it was shown that NCR induced bacteroid enlargement also occurs in nodules of some *Aeschynomene* species that belong to the dalbergioids clade (Fig.1) (Guefrachi et al., 2015; Czernic et al., 2015). As increased bacteroid size occurs rather frequently within the Papilionoideae (e.g. *Tephrosia virginianum* of the milletioids and *Cytisus scoparius* of the genistoids clade (Fig. 1)) (Oono et al., 2010a), it seems possible that NCR triggered bacteroid enlargement evolved multiple times within the Papilionoideae subfamily, or represents an ancestral form which has been lost several times. To obtain more insight in this, we decided to study bacteroid differentiation in the indigoferoids clade. This clade has a crucial phylogenetic position (Fig. 1), as it represents an early branching lineage of the “indigoferoids/milletioids” clade (Lavin et al., 2005; Wojciechowski et al., 2004).

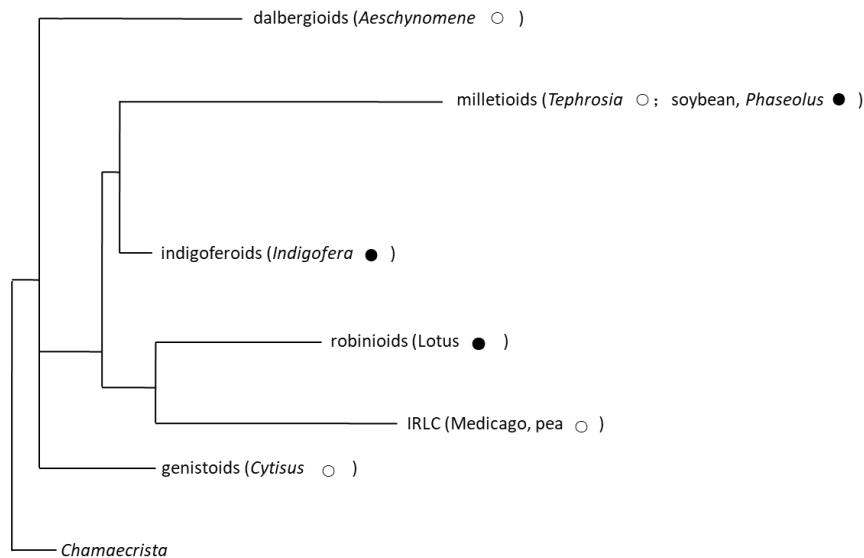


Fig. 1 Clades of the Papilionoideae subfamily. Between brackets names of genera are mentioned and it is indicated which bacteroid morphotype they make. E-/S-morphotype (○) and U-morphotype (●).

In legumes, intracellular bacteroids are in general hosted in a membrane compartment of the host and this transient organelle is named symbiosomes (Roth and Stacey, 1989). Based on their morphology, three distinct bacteroid/symbiosome morphotypes have been described (Oono et al., 2010a; Bonaldi et al., 2011; Czernic et al., 2015). These are elongated bacteroids (E-morphotype), spherical bacteroids (S-morphotype) and unmodified bacteroids (U-morphotype). The latter have a morphology similar to that of free-living bacteria. They occur for example in Lotus, soybean, and *Phaseolus*, and their bacteroids have a DNA content of 1C (Mergaert et al., 2006).

The E-morphotype is very common in the IRLC clade and is correlated with endoreduplication and terminal differentiation. This was first described for bacteroids of, among others, pea nodules (Bisseling et al., 1977), and is now studied in most detail in Medicago (*M. truncatula*) (Mergaert et al., 2003; Mergaert et al., 2006; Wang et al., 2010).

E-morphotype formation has been shown to be correlated with NCR triggered endoreduplication in all studied species. In Medicago, bacteroids increase 5 to 10-fold in size and their DNA content can be as high as 24C (Mergaert et al., 2006). Also the S-morphotype formed in *Aeschynomene* nodules correlates with endoreduplication and on average the bacteroid DNA content is 7C (Guefrachi et al., 2015; Czernic et al., 2015).

Medicago has the best studied NCR family, containing more than 600 genes (Mergaert et al., 2003; Young et al., 2011). It has been shown that these genes are expressed in different regions of the nodule and so are involved in different stages of bacteroid development. Although Medicago has numerous NCR genes, individual NCRs (e.g. *MtNCR169*, *MtNCR211*) are essential for full bacteroid development, as mutations in these NCR genes results in incomplete bacteroid development (Horvath et al., 2015; Kim et al., 2015). Further, the introduction of the *MtNCR035* gene into Lotus is sufficient to induce the formation of enlarged bacteroids (Van de Velde et al., 2010).

MtNCR247 peptide has been reported to penetrate free living *S. meliloti*. It interacts with FtsZ which is required for septum formation and this might contribute to the arrest of cell division (Farkas et al., 2014; Penterman et al., 2014). In free-living *S. meliloti* cells, MtNCR035 initially localizes at the cell envelope and later intracellularly at the bacterial division plane, which is in agreement with a role of MtNCR035 in inhibition of cytokinesis (Van de Velde et al., 2010).

The overall amino acid composition of NCRs is highly diverse, however, they have some common characteristics (Mergaert et al., 2003; Czernic et al., 2015). The mature peptides are relatively short, in general 30-55 amino acids long, and contain 4 or 6 conserved cysteines. Further, all NCR genes are specifically expressed in nodules.

NCRs are related to antimicrobial peptides that are named defensins. Plant defensins are found ubiquitously throughout the plant kingdom and can have for example antibacterial or antifungal activity. In general, mature defensins are 45-70 amino acids long. Their overall amino acid composition also is highly diverse, but they have 8 cysteines. In *Medicago*, most defensin genes have a tissue or organ specific expression, some are, like *NCRs*, specifically expressed in nodules (Maroti et al., 2015).

Other IRLC clade species (e.g. *Glycyrrhiza uralensis*, *Oxytropis lamberti*, *Astragalus canadensis*, *Onobrychis viciifolia*, *Galega orientalis*, *Ononis spinosa*, *Cicer arietinum*, *Pisum sativum*, and *Medicago sativa*) also have NCR families, but in general they are smaller than the *Medicago* NCR family. For example, 353 NCR transcripts were identified in pea and *C. arietinum* has 63 putative NCR genes (Alves-Carvalho et al., 2015; Montiel et al., 2016; Montiel et al., 2017). NCRs in all the studied IRLC clade species share conserved cysteine spacing pattern, containing 4 or 6 cysteines (Fig. 3).

Also outside the IRLC clade bacteroid endoreduplication is correlated with NCR peptides made by the host. For example, *Lotus* and soybean that have U-morphotype bacteroids (1C), do not have NCR genes (Mergaert et al., 2003). In contrast, *Aeschynomene* species have NCR-like families with about 40-80 members (Czernic et al., 2015; Guefrachi et al., 2015). *Aeschynomene* NCR-like peptides are short and the overall amino acid composition also is highly diverse. They are classified into two groups. One group has 8 conserved cysteines giving rise to a typical defensin signature, the other group has 6 cysteines (Czernic et al., 2015). However, the lengths of the loops in between cysteines are highly diverse.

The processing of NCR peptides includes removal of the signal peptide by the signal peptidase complex, of which DNF1 is a subunit (Wang et al., 2010). *Medicago* has a nodule specific *MtDNF1* and a non-symbiotic *MtDNF1-like* gene. A loss of function of *MtDNF1* leads to non-fixing nodules in which bacteria have a U-morphotype. In these nodules, the unprocessed NCRs are not targeted to the symbiosomes. *Aeschynomene* only has a single *DNF1(-like)* gene. Knock down of *Aeschynomene DNF1* resulted in defects in bacteroid differentiation (Czernic et al., 2015). However, as *Aeschynomene* has a single *DNF1(-like)* gene, this knock down may have also affected the non-symbiotic functions of *DNF1-like*, which could have resulted in secondary effects on bacteroid development.

In this study, we show that *I. argentea* forms elongated and branched bacteroids. Bacteroid size and genomic amplification were studied by flow cytometry. This showed that the average increase in size is 7 fold, whereas their average genome size is 3C. Nodule-specific NCR-like peptides were identified from the *I. argentea de novo* transcriptome assembly. Further, in general, >95% of bacteroids are terminally differentiation. It is discussed whether the NCRs are involved in the induction of this terminal differentiation.

Results

Elongated and branched bacteroids in *I. argentea* nodules

We studied the morphology of bacteroids in *I. argentea* nodules. Three months after inoculation with *Bradyrhizobium elkanii* (SA281), bacteroid morphology was studied by light microscopy. This showed that most bacteroids from *I. argentea* nodules are substantially enlarged and several have a branched Y-shape (E-morphotype) (Fig. 2b), whereas free-living *B. elkanii* bacteria are short rods (U-morphotype) (Fig. 2a).

To quantify the increase in size and DNA content, bacteroids were analysed by flow cytometry. This showed that bacteroids have on average a 7-fold enlarged size and their DNA content peaks at 3C (Fig. 2c, d).

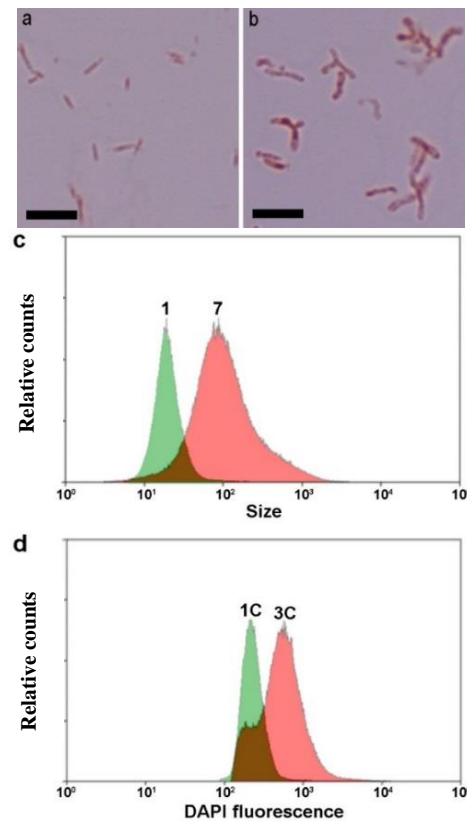


Fig 2. Properties of free-living *B. elkanii* (SA281) bacteria and bacteroids isolated from *I. argentea* nodules. Morphology of free-living bacteria (a) and bacteroids isolated from nodules (b) stained with 0.1% Safranin O; (c) Size of bacteria and bacteroids measured by flow cytometry; (d) DNA content of DAPI-stained bacteria and bacteroids measured by flow cytometry. Green, free-living bacteria; red, bacteroids isolated from nodules. Bars: 5µm.

The viability of bacteroids from five *I. argentea* nodules was measured independently as described in methods. In four of these nodules, > 95% of the bacteroids lost the ability to divide. In one nodule, about 25% cells were able to form colonies.

As *I. argentea* bacteroids have an increased size and DNA content and appear to be terminally differentiated, we studied whether *I. argentea* has NCR genes.

I. argentea NCR peptides

To identify *I. argentea* NCR peptides, a *de novo* transcriptome was constructed using RNA-seq data from nodules, roots, leaves, flowers and pods (chapter 2). Putative NCR-like peptides

a	1	10	20	30	40	50	60
IaNCR1	MEGRK	VIGMLMVLITIKSLV	MLVVADTASTKVTSPSITKTFQS--IEKSS-VPSTGTSDQ				
IaNCR2	MKGKEVIGILMVIAIIESL	VMLVVAD--TKVTSPPI	TKTSQS--MEKSSVVP	SI	VTSDQ		
IaNCR3	MEGK-VIGMLTFMLIVAAY-KITAGDLAPKGI	FSTQISPSTKSQLIEDST----	IKFGAE				
IaNCR1	AHKLGGG-PHCPKVCQQQC	SGSVIG----	CFD	ACMNRC	MFMP	PGFNIRFNI	FHDIGSD
IaNCR2	AHKLGKSVPC	CLRACEQQC	RG	SVDP----	CFETCMSRC	IMMSGSFNMPF	GSIFHHIG-D
IaNCR3	LRDFG----	ECTGLCQMHQ	MDANNPADFPV	CFNNCLGHQ	I	INSPK-----	GT-
IaNCR1	GATNSISMKPN	SQDEAVFEWPRNAVPWPPTFERKG					
IaNCR2	-A-NEVEGKNN	-----					
IaNCR3	--MDNIH	-----					
MsDef2.1	MEKKS	SIAGL-CFLFLILFVAQEIVVTEARTCEHLADTYRGPCF	---	TDAS	CDDHCKNKAH		
MsDef3.2	MEKKS	LAAL-CFLFLVLFVAQEIVVTEAKTCENLADTYRGPCF	---	TNGA	CDDHCKNKEH		
MtDef2.1	MEKKS	SIAGL-CLLFLVLFVAQEIVVTEARTCEHLADTYRGPCFTEG	---	SCDDH	CKNKAH		
PeaDRR39	MEKKS	LAALSFLLLVLFVAQEIVVTEANTCEHLADTYRGVCF	TNA---	SCDDH	CKNKAH		
IaNCR4	MERKS	LAGL-CFELIVFLAPQELIV-KAEKCEKPSILFPEG	CIGEVGKGNCEYVCKIGEG				
MsDef2.1	LISGT	CHR-LQCFCTQNC					
MsDef3.2	LVSGR	CRDDFRCWCTRNC					
MtDef2.1	LISGT	CHN-FQCFCTQNC					
PeaDRR39	LISGT	CHD-WKCFCTQNC					
IaNCR4	LLGGI	CKEHQKCFCA--C					
b	Defensin-like motif: SP-Xn-C-Xn-C-Xn-C-X3-C-Xn-C-Xn-C-X1-C-X3-C-Xn						
	Aeschynomene NCR-motif 2: SP-Xn-C-Xn-C-Xn-C-X3-C-Xn-C-X4-C-X1-C-X3-C-Xn						
	Aeschynomene NCR-motif 1: SP-Xn- C-Xn-C-X3-C-Xn-C-X4-C-X1-C-Xn						
	Indigofera NCR-motif a: SP-Xn-C-Xn-C-Xn-C-X3-C-Xn-C-Xn-C-X1-C-X1-C						
	Indigofera NCR-motif b: SP-Xn- [C-X3-C-X3-C-Xn][C-X3-C-X3-C-Xn]						
	IRLC clade NCR-motifs: SP-Xn- C-X5-C-Xn-C-Xn-C-X4-C-X1-C-Xn						
	SP-Xn- C-X5-C-Xn- C-X4-C-Xn						

Fig. 3 Alignment and cysteine signatures of *I. argentea* NCR peptides. (a) The alignments were made using MEGA5 and adjusted manually; (b) Comparison of the cysteine pattern of IaNCR peptides with those of IRLC clade species, *Aeschynomene* species, and defensin-like peptides.

were selected in the following way. The transcripts are nodule specific (with more than 20 reads mapped to the nodule transcriptome and with less than 5 reads mapped to the other samples). Further, encoded proteins are less than 120 amino acids long, contain 4 or more cysteines in their mature peptide and they have a signal peptide (<http://www.cbs.dtu.dk/services/SignalP/>) (Cut-off threshold = 0.5). This in silico analysis identified 3 putative *I. argentea* NCR peptides. By performing BLASTp searches against the same transcriptome data, using these 3 candidate peptides, 1 more putative NCR peptide was identified. These peptides were named IaNCR1, IaNCR2, IaNCR3, and IaNCR4 (Fig. 3a).

The lengths of the 4 peptides were between 48 and 122 amino acids long (Fig. 3a). IaNCR1, IaNCR2, and IaNCR3 peptides contain 6 cysteine residues, which form two C-X3-C-X3-C domains (shown in boxes in Fig. 3a, b). The signal peptides (underlined) are located at the N-terminus. Further, IaNCR1 and IaNCR2 are 63% identical. The two C-X3-C-X3-C domains are unique in IaNCR peptides, and absent in all known IRLC clade NCRs and *Aeschynomene* NCR peptides. One C-X3-C-X3-C domain is present in some *Medicago* defensins (e.g. MtDef1.1, MsDef1.2) (Hanks et al., 2005).

IaNCR4 peptide has 8 cysteine residues. This cysteine pattern is very similar to that of defensins in for example *Medicago* and pea (MsDef2.1, MsDef3.2, MtDef2.1, and PsDRR39), and is conserved in all studied plant defensins (Van der Weerden, 2013). Their overall homology, however, is low (~ 30%) (Fig. 3a).

The IaNCR4 cysteine pattern is named Indigofera NCR-motif a, and the IaNCR1/2/3 pattern Indigofera NCR-motif b (Fig. 3b).

To study the biological function of IaNCR peptides, we firstly checked the occurrence of the *IaNCR* transcripts in our RNA-seq data and used RT-qPCR to determine expression levels during nodule development.

Expression of *IaNCR* genes during nodule development

Based on the transcriptome analysis, these 4 putative IaNCR genes are expressed in nodules, and transcripts are not detectable in uninfected root, leaf, flower, and pod (Fig. 4a). *IaNCR1* is more than 60-fold higher expressed than the other 3 *IaNCR* genes in 4 weeks old nodule. To confirm the nodule specific expression and to determine the expression level during nodule development of these 4 putative *IaNCR* genes, we performed RT-qPCR analyses.

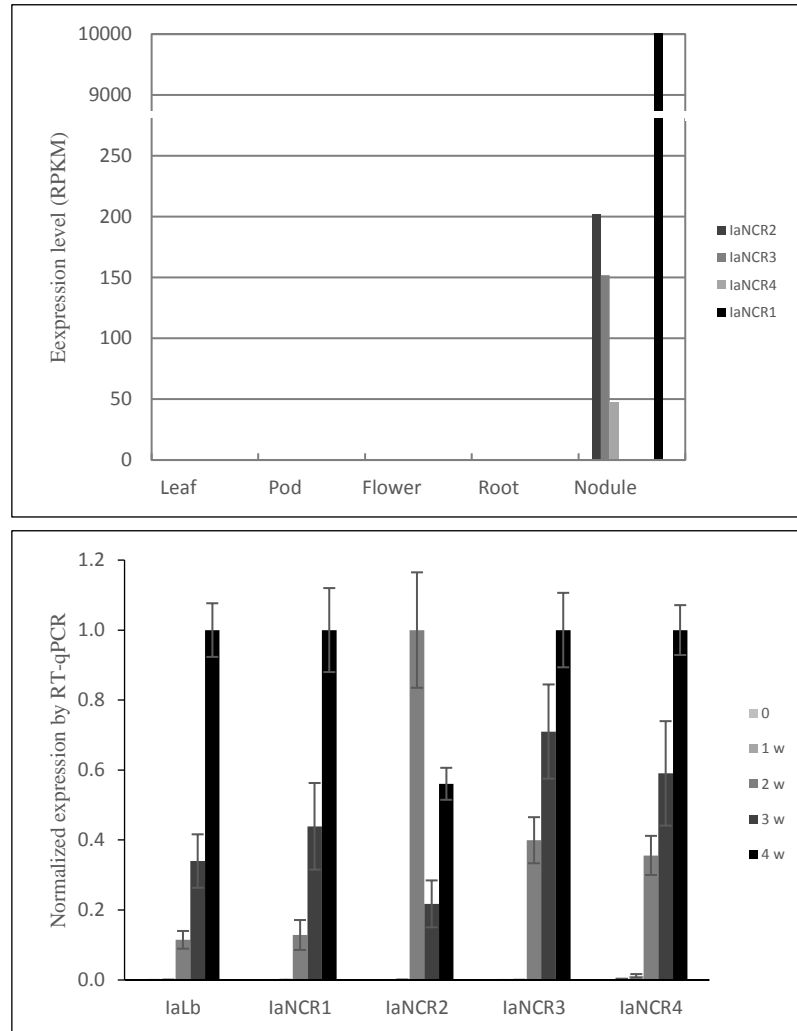


Fig. 4 (a) Expression levels of *IaNCR*-like genes in nodules and non-symbiotic organs based on RNAseq analysis, (b) The relative expression at different nodule stages by RT-qPCR analyses, *IaLb* as a control of nodule development. Error bar: Standard error of the mean, ± 2 SEMs.

RNA was isolated from uninfected roots (time 0) and inoculated roots/nodules at different time points (1, 2, 3, and 4 weeks post inoculation). Four replicates for each time point were collected. *Leghemoglobin* (*IaLb*) mRNA was used as control. *IaLb* transcript started to accumulate at 1 week post inoculation after which it gradually increased. The 4 *IaNCR* transcripts were undetectable in uninfected roots. They started to accumulate at 1 week albeit at a rather low level (Fig. 4b). *IaNCR1*, *IaNCR3* and *IaNCR4* expression increased gradually

at 2, 3, and 4 weeks (Fig. 4b). *IaNCR2* had reached the highest level at 2w after which it remained more or less constant (Fig. 4b).

To further study the role of the IaNCR peptides in bacterial differentiation, we studied first whether *I. argentea* has a nodule specific *DNF1* gene (Van de Veld et al., 2010; Wang et al., 2010).

Identification of *IaDNF1-like* genes

A BLASTp search against the *I. argentea* transcriptome using MtDNF1, identified two highly homologous proteins. These two proteins are both 167 amino acids long with 5 amino acids differences indicated in bold (Fig. 5a). They are 85.1% identical to MtDNF1, 89.9% to MtDNF1-like and 92.9% to AeDNF1, respectively (Fig. 5a) (Czernic et al., 2015). Subsequently, we addressed the question whether one of the isoforms is nodule specifically expressed.

To answer this question, we compared their expression levels in different organs using the RNA-seq data. This showed that the two *I. argentea* *DNF1(-like)* transcripts have a similar expression level in the different organs (Fig. 5b). Both transcripts occur in nodules at a similar level. Based on this expression pattern, we named the two genes *IaDNF1-like1* and *IaDNF1-like2*.

Discussion

Here we show that *B. elkanii* differentiates into E-morphotype bacteroids in *I. argentea* nodules and this is correlated with endoreduplication. Most of the bacteroids lost their viability. This terminal differentiation of bacteroids is accompanied with the expression of several *IaNCR* genes. The encoded IaNCR peptides have a different cysteine signature compared to their counterparts in the IRLC clade or in *Aeschynomene* species. This suggests that NCRs evolved independently in 3 clades of the Papilionoideae subfamily.

B. elkanii forms E-morphotype bacteroids in *I. argentea*, resulting in an about 7-fold increase in size. This increase in size correlates with bacteroid DNA content of an average 3C (Fig. 2). E-morphotype bacteroids in Medicago increase 5 to 10-fold in size and their DNA content can be as high as 24C (Mergaert et al., 2006). DNA content of *Aeschynomene* E- and S-morphotype bacteroids peaks at 7C and 16C, however their size enlargement has not been described (Guefrachi et al., 2015; Czernic et al., 2015). The enlargement of *Bradyrhizobium* bacteroids in *I. argentea* and *Sinorhizobium* bacteroids in Medicago are rather similar. However, the polyploidy level in *Sinorhizobium* bacteroids is much higher than in *Bradyrhizobium* bacteroids in nodules of *I. argentea* and *Aeschynomene*. Therefore, growth of the *Bradyrhizobium* bacteroids in *I. argentea* nodules seems stimulated by lower endoreduplication levels.

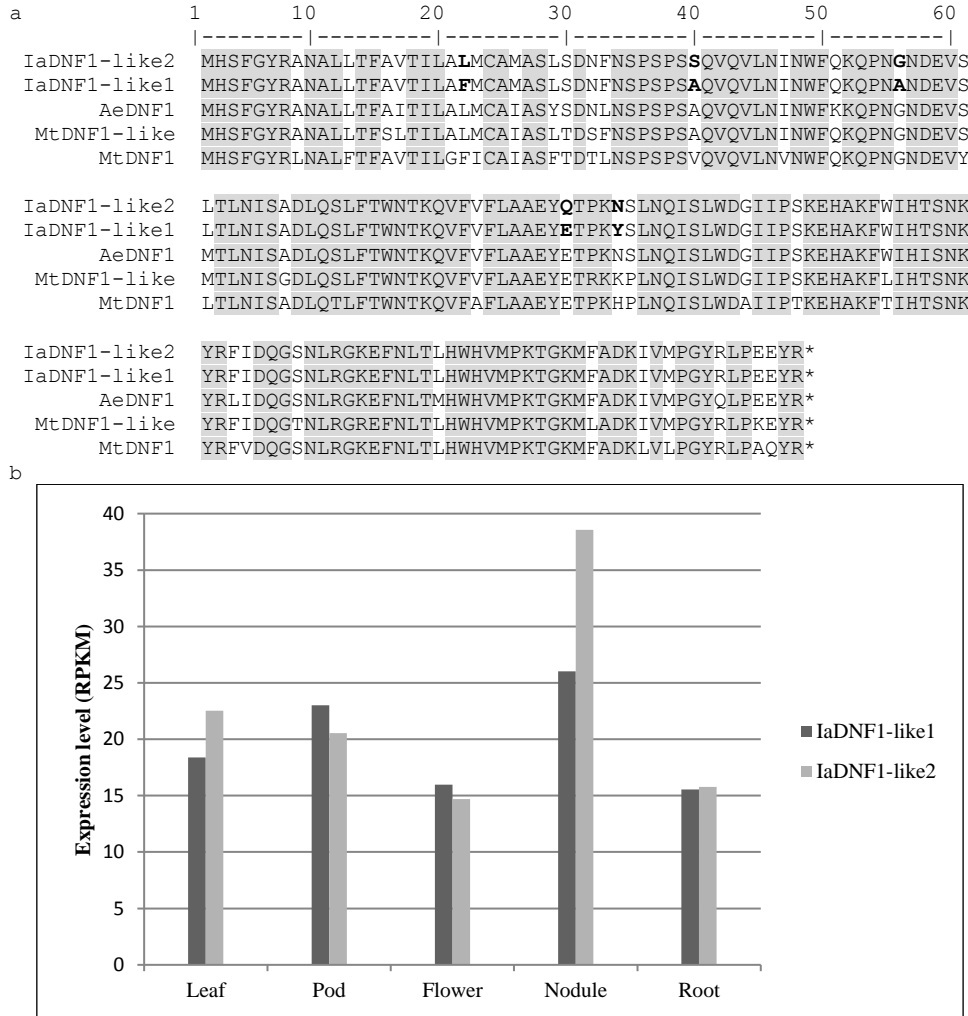


Fig. 5 Identification of *IaDNF1-like1* and *IaDNF1-like2* genes. (a) Alignment of *IaDNF1-like1*, *IaDNF1-like2*, *AeDNF1-like*, *MtDNF1*, and *MtDNF1-like*, the 5 different amino acids of *IaDNF1-like1* and *IaDNF1-like2* are highlight in bold. (b) The expression level of *IaDNF1-like1* and *IaDNF1-like2*. Both are expressed in nodules and non-symbiotic organs.

In four *I. argentea* nodules, >95% of the bacteroids were unable to form colonies on agar plates suggesting that they are terminally differentiated. This was similar to that observed for *S. meliloti* bacteroids from *Medicago* nodules (99%; Mergaert et. al, 2006) and *B. sp.* strain ORS285 bacteroids from *A. indica* nodules (98%; Czernic et. al, 2015). In one nodule, a relatively high percentage of viable cells (25%) was observed we assume that this is can be the result from a secondary cluster of dividing cells that still contain U-morphotype

bacteroids (Chapter 4). This higher number of viable bacteroids is similar to that observed in *A. afraspera* with 22% viable cells. It was speculated that this was caused by the presence of a cluster of cells near the surface of the root in which the bacteria remained unmodified (Bonaldi et al., 2011; Czernic et al., 2015).

The four *IaNCR* genes might be involved in different stages of bacteroid differentiation. *IaNCR2* mRNA has reached a maximal level before *Lb* mRNA, whereas the other 3 have a similar expression pattern as *LB*. Therefore, we expect that *IaNCR2* is involved in earlier stages of development than the other 3 *IaNCRs*.

To study whether the *IaNCR* peptides induce bacteroid elongation, the following strategies could be followed. It can be studied whether transformation of *IaNCR* genes into *Lotus* affects the bacteroid phenotype as in the case of *MtNCR035* (Van de Velde et al., 2010). Further, RNA interference of *IaDNF1-like* genes can be done in *I. argentea* as a hairy root transformation protocol has been developed (Chapter 3). However, The resulting bacteroid phenotype might not be due to unprocessed *IaNCR* peptides but a secondary effect as also non-symbiotic processes will be affected (Czernic et al., 2015).

I. argentea NCR peptides have a low level of homology but the cysteine spacing pattern is conserved. However, this pattern is different from that of NCR peptides identified in IRLC clade and *Aeschynomene* spp. The cysteine spacing pattern of NCR peptides in all studied IRLC clade species (e.g. *G. uralensis*, *O. lamberti*, *A. canadensis*, *O. viciifolia*, *G. orientalis*, *O. spinosa*, *C. arietinum*, *P. sativum*, *M. sativa*, and *M. truncatula*) is similar (Fig. 3b). This indicates that NCRs with this cysteine pattern were already present in their oldest common ancestor (Mergaert et al., 2003). *Aeschynomene* NCR peptides contain cysteine domains with a different spacing pattern (Fig. 3b), and therefore it has been postulated that they evolved independently from those of the IRLC clade (Czernic et al., 2015). The indigoferoids clade has a crucial phylogenetic position (Fig. 1), as it represents an early branching lineage of the “indigoferoids/milletioids” clade (Lavin et al., 2005; Wojciechowski et al., 2004). Most of the studied species in milletioids contain unmodified bacteroids and their genomes do not contain *NCRs*. In *I. argentea*, *IaNCR1*, *IaNCR2*, and *IaNCR3* peptides contain two C-X3-C-X3-C domains, which are not present in the IRLC clade and *Aeschynomene* species (Fig. 3b). It seems that NCRs driven E- and S-morphotype bacteroid differentiation evolved three times independently in these three clades. The distinct cysteine patterns of NCRs in these three clade indicates that they evolved from defensins, but from different ones. This explains the rapid NCRs evolutionary amplification and diversification in *Medicago* or IRLC clade. The convergent evolution of NCRs in Papilionoideae subfamily suggest that this is beneficial for host plant. For example enlarged bacteroids have higher nitrogen fixation efficiency and so host plants grow better (Oono et al., 2010b).

Further the cysteine rich peptides might be part of a strategy to control large populations of intracellular bacteria. NCRs most likely evolved from defensins. These occur wide spread in the plant kingdom. Therefore, it is of interest to determine whether NCRs also evolved in non-legume nodulating plant species, for example in actinorhizal species. In the nodule transcriptome of the actinorhizal plant *Datisca glomerata*, *DgCRP1* encodes a basic cysteine-rich peptide containing C-X3-C-X4-C-X10-C-X3-C-X3-C (Demina et al., 2013; Carro et al., 2014). This cysteine pattern is similar to that in *I. argentea*. The occurrence of this NCR like peptide in *Datisca* nodule supports the conclusion that the NCR system is an advantage and therefore it evolved several times most likely even outside the legume family.

Materials and methods

Plant material and bacterial strains

I. argentea plants were grown under controlled conditions as follow: 11h/13h (day/night), 35 °C/25 °C (day/night), and 60% humidity. *Bradyrhizobium elkanii* (SA281) was grown at 28 °C in yeast mannitol medium (YEM) for 4-5 days. Seed germination and root inoculation were performed as described (Chapter 2). Nodules were harvest 10 day, 2 weeks, 3 weeks, 4 weeks, 5 weeks, and 3 months after inoculation.

Bacteroid characterization and viability determination

Bacteroid extraction was performed as described (Mergaert et al., 2006). Bacteroids and free-living bacteria were stained with 0.1% Safranin O and analysed with an AU5500B microscope equipped with a DFC425c camera (both Leica). Bacteroid size and DNA content were measured by ASTRIOS Beckman Coulter machine as described (McRae et al., 1989; Mergaert et al., 2006).

Bacteroids were extracted from five nodules independently following the methods. Together with free-living *B. elkanii* (SA281), series dilutions were prepared and plated on the YEM medium. After growing at 28 °C for five days, colonies formed were counted.

IaNCR transcripts identification

RNA isolation, sequencing, and *de novo* transcriptome analysis were performed as described (Chapter 2). The proteins were predicted based on the longest ORFs using Transdecoder (Haas et al., 2013) with a minimum length of 30 amino acids.

IaDNF1-like peptide sequence and phylogenetic analyses

The IaDNF1-like proteins were identified from the transcriptome database using MtDNF1 protein (XP_003599103) as query sequence. This BLASTp search allowed the identification

of two candidate proteins, *Ia_c55598_g1_i1* and *Ia_c55598_g2_i1*, both were of 167 amino acids.

Real Time quantitative PCR expression analysis

Total RNA was extracted, using phenol-chloroform extraction, from uninfected roots or roots with nodules at different time points after inoculation (Yang et al., 2008). Four biological replicates were taken for each time point. After DNase I (Invitrogen) treatment, the integrity of total RNA was tested by gel electrophoresis. cDNA was synthesized from 400ng total RNA per sample using iScript cDNA synthesis kit (Bio-Rad, Hercules, USA). Real-time quantitative PCR reactions were set up in a 20µl system with 2×iQ SYBR Green Super-mix (Bio-Rad, Hercules, USA) and the CFX96 Manager Real-time PCR detecting system according to the manufacturer's manuals. The primers used for quantitative PCR were designed using the qPCR settings of Primer3Plus (Untergasser et al., 2007) and listed in Supplemental Table S1. Reactions were performed in triplicate. Relative expression values were normalized to *IaPTB*, *IaSAND*, and *IaUNK1* reference genes (Kakar et al., 2008; Zhu, 2013), which are stable expressed in different organs (Fig. S1).

References

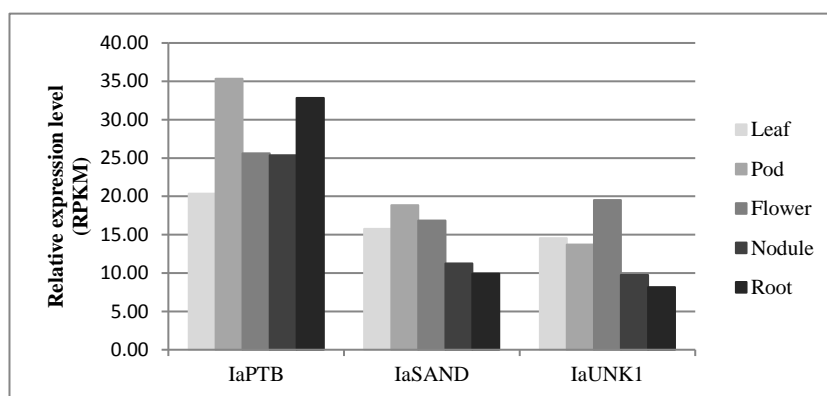
- Alves-Carvalho, S., Aubert, G., Carrere, S., Cruaud, C., Brochot, A.L., Jacquin, F., Klein, A., Martin, C., Boucherot, K., Kreplak, J., Da Silva, C., Moreau, S., Gamas, P., Wincker, P., Gouzy, J., and Burstin, J. (2015). Full-length de novo assembly of RNA-seq data in pea (*Pisum sativum* L.) provides a gene expression atlas and gives insights into root nodulation in this species. *Plant Journal* 84(1): 1-19.
- Bisseling, T., Vandenbos, R.C., Van Kammen, A., Vanderploeg, M., Vanduijn, P., and Houwers, A. (1977). Cytofluorometrical determination of the DNA contents of bacteroids and corresponding. *Journal of General Microbiology* 101: 79-84.
- Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., Goormachtig, S., and Giraud, E. (2011). Nodulation of *Aeschynomene afraspera* and *A. indica* by photosynthetic *Bradyrhizobium* Sp. Strain ORS285: The Nod-dependent versus the Nod-independent symbiotic interaction. *Molecular Plant-Microbe Interactions* 24(11): 1359-1371.
- Carro, L., Pujic, P., Alloisio, N., Fournier, P., Boubakri, H., Hay, A.E., Poly, F., Francois, P., Hoher, V., Mergaert, P., Balmand, S., Rey, M., Heddi, A., and Normand, P. (2015). *Alnus* peptides modify membrane porosity and induce the release of nitrogen-rich metabolites from nitrogen-fixing *Frankia*. *ISME J* 9(8): 1723-1733.
- Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., Pierre, O., Fardoux, J., Chaintreuil, C., Nguyen, P., Gressent, F., Da Silva, C., Poulain, J., Wincker, P., Rofidal, V., Hem, S., Barriere, Q., Arrighi, J.F., Mergaert, P., and Giraud, E. (2015). Convergent evolution of endosymbiont differentiation in dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by Nodule-Specific Cysteine-Rich peptides. *Plant Physiology* 169(2): 1254-1265.
- Demina, I.V., Persson, T., Santos, P., Plaszczyc, M., and Pawlowski, K. (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): e72442.
- Farkas, A., Maroti, G., Durgo, H., Gyorgypal, Z., Lima, R.M., Medzihradsky, K.F., Kereszt, A., Mergaert, P., and Kondorosi, E. (2014). *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proceedings of the National Academy of Sciences of the United States of America* 111(14): 5183-5188.
- Guefrachi, I., Pierre, O., Timchenko, T., Alunni, B., Barriere, Q., Czernic, P., Villacija-Aguilar, J.A., Verly, C., Bourge, M., Fardoux, J., Mars, M., Kondorosi, E., Giraud, E., and Mergaert, P. (2015). *Bradyrhizobium* BclA is a peptide transporter required for bacterial differentiation in symbiosis with *Aeschynomene* legumes. *Molecular Plant-Microbe Interactions* 28(11): 1155-1166.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., Macmanes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N., and Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8(8): 1494-1512.
- Hanks, J.N., Snyder, A.K., Graham, M.A., Shah, R.K., Blaylock, L.A., Harrison, M.J., and Shah, D.M. (2005). Defensin gene family in *Medicago truncatula*: structure, expression and induction by signal molecules. *Plant Mol Biol* 58(3): 385-399.
- Horvath, B., Domonkos, A., Kereszt, A., Szucs, A., Abraham, E., Ayaydin, F., Boka, K., Chen, Y.H., Chen, R.J., Murray, J.D., Udvardi, M.K., Kondorosi, E., and Kalo, P. (2015). Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the *Medicago truncatula* dnf7 mutant. *Proceedings of the National Academy of Sciences of the United States of America* 112(49): 15232-15237.
- Kakar, K., Wandrey, M., Czechowski, T., Gaertner, T., Scheible, W.R., Stitt, M., Torres-Jerez, I., Xiao, Y., Redman, J.C., Wu, H.C., Cheung, F., Town, C.D., and Udvardi, M.K. (2008). A community resource for high-throughput quantitative RT-PCR analysis of transcription factor gene expression in *Medicago truncatula*. *Plant Methods* 4: 18.
- Kim, M., Chen, Y., Xi, J., Waters, C., Chen, R., and Wang, D. (2015). An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* 112(49): 15238-15243.
- Lavin, M., Herendeen, P.S., and Wojciechowski, M.F. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic Biology* 54(4): 575-594.
- Maroti, G., Downie, J.A., and Kondorosi, E. (2015). Plant cysteine-rich peptides that inhibit pathogen growth and control rhizobial differentiation in legume nodules. *Current Opinion in Plant Biology* 26: 57-63.

- McRae, D.G., Miller, R.W., Berndt, W.B., and Joy, K. (1989). Transport of C4-dicarboxylates and amino acids by *Rhizobium meliloti* bacteroids. *Molecular Plant-Microbe Interactions* 2: 273–278.
- Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., and Kondorosi, E. (2003). A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiology* 132(1): 161–173.
- Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., and Kondorosi, E. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium-legume* symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* 103(13): 5230–5235.
- Montiel, J., Szucs, A., Boboescu, I.Z., Gherman, V.D., Kondorosi, E., and Kereszt, A. (2016). Terminal bacteroid differentiation is associated with variable morphological changes in legume species belonging to the Inverted Repeat-Lacking Clade. *Molecular Plant-Microbe Interactions* 29(3): 210–219.
- Montiel, J., Downie, J.A., Farkas, A., Bihari, P., Herczeg, R., Balint, B., Mergaert, P., Kereszt, A., and Kondorosi, E. (2017). Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proceedings of the National Academy of Sciences of the United States of America* 114(19): 5041–5046.
- Oono, R., Schmitt, I., Sprent, J.I., and Denison, R.F. (2010a). Multiple evolutionary origins of legume traits leading to extreme rhizobial differentiation. *New Phytologist* 187(2): 508–520.
- Oono, R., and Denison, R.F. (2010b). Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiology* 154(3): 1541–1548.
- Penterman, J., Abo, R.P., De Nisco, N.J., Arnold, M.F.F., Longhi, R., Zanda, M., and Walker, G.C. (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): 3561–3566.
- Roth, L.E., and Stacey, G. (1989). Bacterium release into host cells of nitrogen-fixing soybean nodules: The symbiosome membrane comes from three sources. *European Journal of Cell Biology* 49(1): 13–23.
- Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., and Leunissen, J.A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* 35 (Web Server issue): W71–74.
- Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A., Tiricz, H., Satiat-Jeunemaitre, B., Alunni, B., Bourge, M., Kucho, K., Abe, M., Kereszt, A., Maroti, G., Uchiumi, T., Kondorosi, E., and Mergaert, P. (2010). Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327(5969): 1122–1126.
- Van der Weerden, N.L., and Anderson, M.A. (2013). Plant defensins: Common fold, multiple functions. *Fungal Biology Reviews* 26(4): 121–131.
- Wang, D., Griffiths, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., Bisseling, T., and Long, S.R. (2010). A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* 327(5969): 1126–1129.
- Wang, Q., Yang, S., Liu, J., Terecskei, K., Abraham, E., Gombar, A., Domonkos, A., Szucs, A., Kormoczi, P., Wang, T., Fodor, L., Mao, L., Fei, Z., Kondorosi, E., Kalo, P., Kereszt, A., and Zhu, H. (2017). Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. *Proceedings of the National Academy of Sciences of the United States of America* 114(26): 6854–6859.
- Wojciechowski, M.F., Lavin, M., and Sanderson, M.J. (2004). A phylogeny of legumes (Leguminosae) based on analyses of the plastid matK gene resolves many well-supported subclades within the family. *American Journal of Botany* 91(11): 1846–1862.
- Yang, G., Zhou, R., Tang, T., and Shi, S. (2008). Simple and efficient isolation of high-quality total RNA from *Hibiscus tiliaceus*, a mangrove associate and its relatives. *Prep Biochem Biotechnol* 38(3): 257–264.
- Yang, S., Wang, Q., Fedorova, E., Liu, J., Qin, Q., Zheng, Q., Price, P.A., Pan, H., Wang, D., Griffiths, J.S., Bisseling, T., and Zhu, H. (2017). Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proceedings of the National Academy of Sciences of the United States of America* 114(26): 6848–6853.
- Young, N.D., Debelle, F., Oldroyd, G.E., Geurts, R., Cannon, S.B., Udvardi, M.K., Bénédict, V.A., Mayer, K.F.X., et al. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480(7378): 520–524.
- Zhu, J., Zhang, L., Li, W., Han, S., Yang, W., and Qi, L. (2013). Reference gene selection for quantitative real-time PCR normalization in *Caragana intermedia* under different abiotic stress conditions. *PLoS One* 8(1): e53196.

Supplemental data:

Table S1. List of primers used for qPCR.

Gene Name	Forward primer	Reverse primer
IaNCR1	CCCGGCTTCAACATACGTTTTC	TCAAATGTTGGTGGCCATGG
IaNCR2	GAGTCGTTGCATCATGATGTCC	TTCCTTCCACTTCGTTTGCG
IaNCR3	GGAGAATGCACTGGTTTGTGTC	GGCTGTTGATGATACAGTGACC
IaNCR4	GGATGCATTGGAGAAGTTGGC	TGCCACCAAGTAAACCTTCTCC
IaLb	TGCTTTCAGTGAGCAGCAAG	AGCGGAGAAGAACAAGCAC
IaSAND	TGCTGGCCAACTTTGTGATG	TCTTGGCAAGCAAACAGGAG
IaUNK1	AAAATCATCAGCGCCCTCAG	TGATTGGGCTGCACTTGTG
IaPTB	AAGTACGCGGGAAAACGTGTG	ATGCTTACAAGGCGTGCATC

Fig. S1 Expression pattern of reference genes *IaPTB*, *IaSAND*, and *IaUNK1* in different plant tissues.

Chapter 6

General Discussion

I initiated research on *Indigofera argentea* as this species and its micro-symbiont are heat and drought resistant. Such properties are not very prominent in the model legumes *Medicago* and *Lotus*, and so cannot really be studied in these species. The nodule material that was harvested in the Jizan desert showed that these nodules could be rather large and most likely were the result of several waves of growth. To determine whether this was an adaptation to the desert habitat, I analysed nodule development. This revealed that this “wavy” indeterminate growth is the result from newly formed (secondary) clusters of dividing cells. This finding showed that the classification of *Indigofera* nodules as classical indeterminate nodules was wrong. Therefore, I studied the evolution of determinate and indeterminate nodules within the Papilionoideae subfamily and this is described and discussed in the second part of this Discussion. In the first part of the Discussion, the characterization and genetic diversity of micro-symbiont (rhizobia) from *Indigofera*, as well as the development of a technology platform for *Indigofera* are discussed.

1. *Indigofera argentea*, a future system to study drought tolerance

I. argentea grows as a pioneer in former agricultural areas in Jizan desert of Saudi Arabia. This region is subtropical and has an arid hot climate. To study mechanisms that control adaptation of abiotic desert condition, I selected *I. argentea* as a system for future molecular genetic studies.

In Chapter 2 I showed *I. argentea* has characteristics that will facilitate molecular genetic approaches. *I. argentea* is a diploid self-pollinating legume. It has a relatively small genome (690 Mb) and short generation time. A procedure for *Agrobacterium rhizogenes*-mediated hairy root transformation was developed and this can facilitate functional analyses of genes involved in the nodule symbiosis. For example, genes involved in the formation of secondary clusters of dividing cells in nodules or adaptation to drought. The *I. argentea de novo* transcriptome dataset could be used for such functional analyses.

In chapter 3 Rhizobia were isolated from *I. argentea* “inoculated” with Jizan desert sand. Isolated rhizobia strains were shown to belong to three genera, *Bradyrhizobium*, *Sinorhizobium*, and *Microvirga*. Efficient nodules induced by rhizobia from these three genera and the *Burkholderia* (β -rhizobia) suggest *I. argentea* is a promiscuous legume. The molecular basis for this promiscuity is not known. I hypothesize it is related to the structure of the Nod factor receptors. *Bradyrhizobium elkanii* strain SA281 induces a high number of efficient nodules (Chapter 3) and was selected for further studies.

2. Evolution of legume nodulation

The Leguminosae family is divided into six subfamilies: Caesalpinioideae, Cercidoideae, Detarioideae, Dialioideae, Duparquetioideae, and Papilionoideae (LPWG, 2017; Sprent et al., 2017). Nodulation is uncommon in the Caesalpinioideae, more common in Mimosoideae

and very common in Papilionoideae (Allen & Allen, 1981; Lewis et al., 2005; Sprent, 2007). The Papilionoideae subfamily includes well studied model species (e.g. *Medicago*, *Lotus*) and important crops (e.g. pea, soybean, *Phaseolus*). Nodule forming species that are studied and discussed in this thesis, belong to this Papilionoideae subfamily. In addition, nodule formation of *Chamaecrista* of the Caesalpinioideae subfamily is analysed (in this chapter), as it is a basal legume. Legume nodules are in general composed of a central tissue that harbours (infected) cells where the rhizobia are hosted intracellularly. These intracellular rhizobia are surrounded by a host membrane and are named symbiosomes (Roth and Stacey, 1989). The central tissue is surrounded by peripheral tissues and in one of these, the nodule parenchyma, the nodule vascular bundles are located. This peripheral position of the vasculature is a major difference with nitrogen fixing root nodules that are made by non-legumes as *Parasponia* and actinorhizal plants (e.g. *Alnus*, *Casuarina*), as all these non-legume nodules have a central vascular bundle (Price et al., 1984; Van Ghelue et al., 1996; Pawlowski and Bisseling, 1996).

Some legumes make nodule-like structures that are less complex. An example is *Gleditsia triacanthos*, belonging to Umtiza clade in the Caesalpinioideae subfamily, that makes “swellings” (Fig. 1a) (Fehér and Bokor, 1926; Friesner, 1926; Bryan et al., 1996). These are formed by mitotic activation of root cortical cells (Fig. 1b, c). Whether these swellings are ancestral to the more complex root nodules is not known.

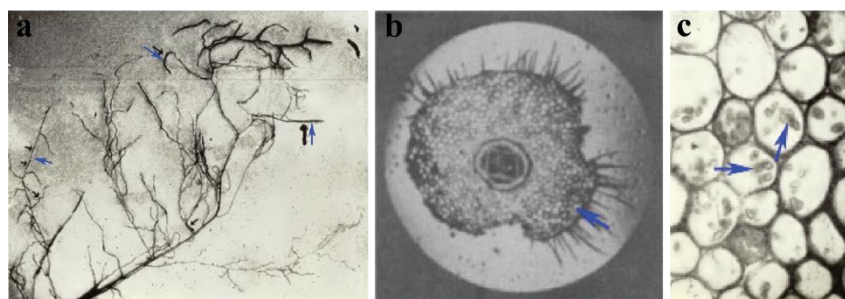


Fig. 1 “Swellings” on *Gleditsia triacanthos* roots according to Friesner (1926), Fehér and Bokor (1926). (a) “Swellings” on the root (arrow); (b) Cross section of a root containing a “swellings” (arrow); (c) Cells harbouring intracellular bacteria (arrow) in the “swellings”.

In this Discussion, I will not discuss the origin of nodulation in legume family, but focus on the evolutionary relationship of the 2 major nodule types formed in the Papilionoideae subfamily; the determinate and indeterminate nodules.

2.1 Determinate and indeterminate nodules

A major difference between determinate and indeterminate nodules is the absence/presence of a persistent meristem. This classification was first used to distinguish the nodules from the

milletioids (e.g. soybean, *Vigna*, *Phaseolus*) and IRLC clade (e.g. pea, *Medicago*, *Vicia*, *Trifolium*). Later it was also used to classify legume nodules from other groups (see for example Fig. 2).

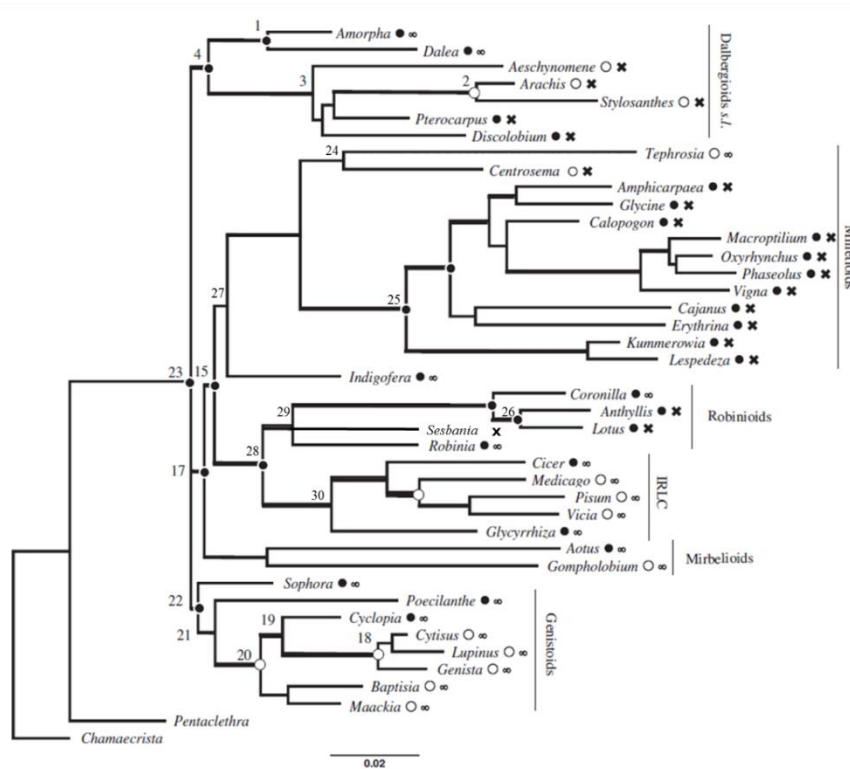


Fig.2 Distribution of indeterminate and determinate nodule type as well as different bacteroid morphotypes of studied species within the Papilionoideae subfamily according to Oono et al. (2010). Indeterminate (∞) and determinate (x) nodule type, E- and S-morphotype (open circle) as well as U-morphotype (closed circle) bacteroid type are indicated next to the genus name.

Indeterminate nodules of the IRLC clade have an apical meristem. This is formed during the primordium stage and it remains active till the nodule senesces. A characteristic of this meristem is that it adds cells to all nodule tissues and further its cells are not infected, i.e. neither infection threads nor intracellular rhizobia occur in these cells. The formation of this meristem is best studied in *Medicago* (Xiao et al., 2014), but it is most likely similar in all IRLC clade species. During primordium formation the inner root cortical cell layers (in *Medicago* layer 4 and 5) and the middle cortical cell layer (layer 3) are mitotically activated. The cells derived from the inner cortex become infected and will ultimately form ~8 proximal cell layers in the (*Medicago*) nodule central tissue. The middle cortical cells are first

mitotically activated when the infection threads (coming from epidermis) has passed/reached this layer. At this stage, the inner cortical layers have already divided several times and have formed multiple cell layers that will be infected. This sequence of events seems important to create a bacterial free meristem at the apex of the primordium/nodule.

The early stages of determinate nodule formation of milletioids species have been best studied in soybean. This was first described in 1958 by Bergersen and Briggs. They studied 1-5 week old nodules and showed that in 1 week old primordia, infection threads are present in the majority of the host cells of the central tissue. They further describe: "These threads traversed dividing cells whose nuclei were in various stages of mitosis and free bacteria were also seen in dividing host cells. Infection threads were difficult to find in cells of 2 weeks or older nodules, although many dividing cells containing bacteria were seen. It seemed therefore, that initially the bacteria were distributed through the dividing cells of the young developing nodule by ramification of infection threads from which the bacteria were released into the cytoplasm of the cells, the bacteria being continuously re-distributed by division of the host cells containing them. Host cell division became less frequent and finally ceased after about 2 weeks....." This process of primordium/nodule development is most likely very similar in other milletioids species.

So a major difference of the IRLC indeterminate nodule and milletioids determinate nodule type is the absence of a non-infected meristem in the determinate nodules. Further, the primordium cells that are penetrated by infection threads and contain released bacteria in their cytoplasm remain mitotically active during determinate nodule formation. However, they stop dividing at an early stage of development and therefore these nodules have a determinate growth type and obtain a spherical shape. In contrast, in IRLC species the nodule primordium cells that are penetrated by an infection threads cease dividing. The meristem is first formed from a middle cortical cell layer when the infection threads have traversed this layer. This meristem remains active throughout the lifespan of the nodule, by which they have an indeterminate growth and obtain an elongated shape. This apical meristem adds cells to all nodule tissues, including the peripheral tissues. In soybean it is shown that the infected cells are transiently mitotically active and in this way support growth of the central tissue. However, how the peripheral tissues grow when the central tissue becomes larger has to my knowledge not been studied in milletioids species.

Interestingly, in the milletioids species the outermost root cortical cells are the first cells that become mitotically activated. These cells are penetrated by infection threads from which bacteria are released. Therefore a more outward cortical cell layer from which a non-infected apical meristem could be formed is not available. I hypothesize that in case the primordium is formed from the outermost cortical layer, an apical meristem composed of non-infected cells cannot be formed.

In Chapter 4 of this thesis I studied the development of *Indigofera* and *Tephrosia* nodules and compared this with the indeterminate and determinate nodule development of the IRLC clade and other millettoids species, respectively.

2.2 *Indigofera* nodule development

Indigofera nodules have previously been classified as indeterminate nodules (Fig. 2). However, we showed that their development differed from that of the indeterminate nodules of the IRLC clade (Chapter 4). A major difference is that, like millettoids nodules, the *Indigofera* primordia are formed from the outermost cortical cell layer and infected primordium cells maintain mitotically active during early stages of nodule development. A meristem composed of non-infected cells is not formed in the primordia. So the early stages of *Indigofera* nodule development are similar to that of the millettoids species. In older *Indigofera* nodules often a cluster of infected dividing cells is present and they are often separated from the rest of the central tissue by layers of uninfected cells, which indicates they originate from mitotically activated nodule boundary or parenchyma cells (Fig. 2 in General Introduction). This is most obvious in two species, *I. argentea* and *I. tinctoria*. In the other three analysed *Indigofera* species from the Tethyan clade, the cluster of dividing cells is located adjacent to infected cells of the central tissue. These secondary clusters of dividing cells are responsible for the indeterminate growth of these nodules. However, the developmental origin of this cluster is different from that of the non-infected meristem of the IRLC species. An additional difference is that it is composed of cells that are infected as they contain infection threads or released bacteria. These cells are most likely in part differentiated and their fate to become the infected cell type is determined. In contrast, the IRLC nodule meristem cells seem undifferentiated and can develop into all nodule cell types.

The difference between millettoids and *Indigofera* nodules seems the ability to form secondary clusters of dividing cells. However, this is not a black and white difference. Iconic determinate soybean nodules can also form secondary clusters of dividing cells that are formed from boundary/parenchyma tissue. However, these have so far only been observed after a stress treatment and their formation seems to depend on the light quality (James et al., 1993). So millettoids and *Indigofera* nodule development are similar and only some gradual/subtle differences distinguish them. These differences concern for example the ease by which secondary clusters of dividing cells can be created, after salt stress or just under lab condition.

2.3 *Tephrosia* nodule development

As the early development of *Indigofera* nodules showed to be very similar to that of the millettoids species, I decided to analyse nodule development of *Tephrosia*, the only millettoids genus that has been reported to make indeterminate nodules (Fig. 2). In the three *Tephrosia* species that have been studied (Chapter 4), primordia are induced in the outermost

cortical cell layers. In line with this, an apical meristem composed of non-infected cells is not formed. Infected primordium cells containing released bacteria remain mitotically active during early stages of development. In general, young nodules are very similar to that of other milletioids and *Indigofera* species. In older nodules, clusters of dividing cells can be present. They are located adjacent to infected cells of the central tissue, which is similar to that in the nodules of three studied *Indigofera* species from the Tethyan clade.

So *Tephrosia* makes nodules that are very similar to that of the other milletioids species. A subtle difference is that clusters of dividing cells are more easily formed like in the *Indigofera* species.

2.4 Indeterminate nodule growth in different clades is the result of convergent evolution

To understand the evolution of nodule formation it is essential to use the underlying developmental processes. As these processes controlling indeterminate growth of *Indigofera* and *Tephrosia*, and IRLC nodules are different, this indeterminate growth behavior appears to be the result of convergent evolution. *Indigofera* and *Tephrosia* nodule development share with determinate nodules of the milletioids species that they lose their primary mitotic activity at an early stage of development. Further, both *Indigofera* and milletioids nodules are derived from the outermost cortical cell layer. In relation to how meristems are formed during IRLC nodule development, this seems a very relevant criterion that distinguishes these nodule types. As discussed above it seems more or less impossible to form an apical meristem composed of cells that are not infected when the primordium is derived from the outermost cortical cell layer. To my knowledge this has also never been described. However, when primordia are formed from more inner located cortical cells, it does not mean that always a non-infected meristem is formed. An example is *Lotus* in which primordia are formed from middle cortical cell layers, but it has a typical determinate nodule development (Szczyglowski et al., 1998).

The meristem of IRLC nodules adds new cells to all nodule tissues: the central tissue as well as the peripheral tissues. In contrast, the clusters of dividing cells in *Indigofera* and *Tephrosia* only seem to add cells to the central tissue, whereas the peripheral tissues grow by divisions in cells of these tissues (Chapter 4). Therefore, I propose that the cells of the IRLC nodule meristem are stem cell-like as they can develop into different cell types, whereas the developmental program that the dividing cells in *Indigofera* and *Tephrosia* will follow is already determined.

The processes underlying indeterminate growth in the IRLC clade, and that of *Indigofera* and milletioids appear to be very different. This implies that they evolved independently. This indicates that there is pressure to evolve nodules with indeterminate growth. The reason for this could for example be, to create new infected cells when the older cells start to senesce

(Chapter 4); to adapt to varying environmental conditions; or to support growth during different growth seasons of a perennial.

To position the development of *Indigofera* and *Tephrosia* nodules in a broader evolutionary perspective, I first provide nodule developmental analyses of some other legume species in the Papilionoideae subfamily.

2.5 Nodule development in the genistoids

Lupin is a rather well studied genus within the Papilionoideae subfamily and is part of the genistoids clade. This clade has a common ancestor (node 23 in Fig. 2) with the dalbergioids, mirbelioids, IRLC, robinoids, *Indigofera* and milletioids clades. Lupin (e.g. white lupin) nodules have been classified as lupinoid nodule, which is considered to be a subtype indeterminate nodule. White lupin nodule primordia are formed by mitotic activation of the outermost cortical cells (Fig. 3a) (Tang et al., 1992; James et al., 1997), and consistent with this, a nodule meristem composed of non-infected cells is not formed at the primordium stage, the infected primordium cells maintain mitotically active for some time (Fig. 3d). This is similar to *Indigofera* and milletioids nodule development. Infection threads are initiated in root hairs and these penetrate the primordium cells and bacteria are released. These dividing cells contain a relatively high number of symbiosomes (Fig. 3d). The infected cells remain mitotically active and the symbiosomes are equally distributed between the daughter cells (Fig. 3d) (González-Sama et al., 2004; Fedorova et al., 2007). This is similar to soybean, *Tephrosia*, and *Indigofera* in which primordium cells are mitotically active with infection threads and in several cases also after release of rhizobia from infection threads (Chapter 4) (Bergersen and Briggs, 1958). However, the number of symbiosomes in the dividing infected cells of Lupin nodules is higher than in *Indigofera* and milletioids species. This is probably the reason that this was considered a specific property of lupinoid nodules. However, the property that dividing cells can be infected cells is wide spread and probably only lacking in IRLC nodule meristem cells.

Mitotic activity in Lupin nodules stops at an early stage, no dividing cells are present in the central tissue of nodules 4 weeks after inoculation (Fig. 3b). In nodules harvested at 5 weeks, clusters of dividing infected cells (Fig. 3c) are present, these are most likely secondary formed clusters like in for example *Indigofera*. However, these dividing cells are located at the basal-lateral side of lupinoid nodules (Fig. 3c, 3d). So the position of dividing cells in Lupin nodules differs from that of *Indigofera* and *Tephrosia*. The basal-lateral dividing infected cells result in broadening of the connection of the nodule to the root, resulting in a characteristic lupinoid nodule morphology (Gonzalez-Sama et al., 2004). In for example *Medicago* nodules (primordia), it has been shown that cell division correlates with a high auxin response level (Franssen et al., 2015). This is also the case for the clusters of dividing infected cells in *Indigofera* (Chapter 4). So the early stages of Lupin nodule development are

very similar to that of the millettoids and *Indigofera*, and different from the indeterminate nodules of the IRLC clade as a non-infected persistent meristem is lacking. Lupin shares with *Indigofera* and millettoids that secondary clusters of dividing cells can be formed. However, the position of these cells is different, basal lateral versus nodule apex.

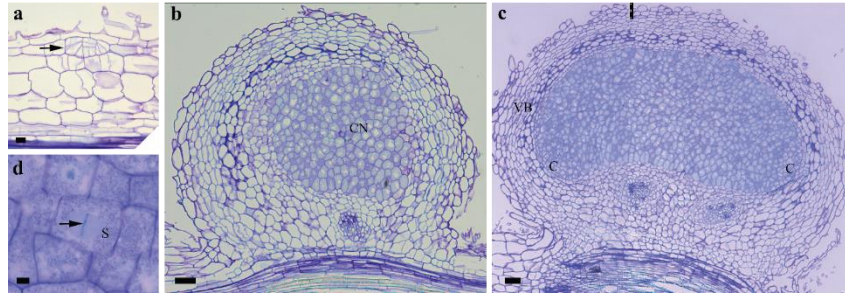


Fig. 3 Longitudinal section of primordia and nodules formed on white lupin. (a) Primordia are induced in the outer most cortical cells layers; (b) Determinate nodules formed; (c) Cluster of dividing cells located at the basal-lateral bottom of nodule; (d) Dividing cell at metaphase with bacteria inside. VB, vascular bundle; C, cluster of dividing cells; CN, central zone; S: symbiosome; Scale bars: a, 25µm; b-c, 75µm; d, 5µm.

To study whether this Lupin nodule developmental process is common in the genistoids, I analysed three more species, *Sophora flavescens*, *Ulex europaeus*, and *Cytisus scoparius*. *Sophora* represents an early branching lineage within the genistoids clade. *Genista* and *Cytisus* are the two genera most closely related to Lupin (Fig. 2). Nodule development analyses showed that primordia were also formed from the outermost cortical cell layer in *Sophora* (Fig. 4a), *Ulex* (Fig. 5a), as well as *Cytisus* (Fig. 6a). Infected primordium cells maintain mitotically active and a non-infected meristem is not formed at primordium stage (Fig. 4b; 5b; 6b). At later stages, dividing infected cells with bacteria are present at the

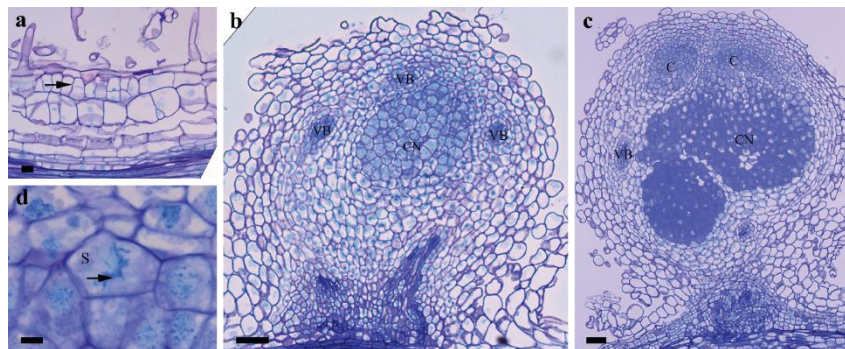


Fig. 4 Longitudinal section of primordia and nodules formed on *S. flavescens*. (a) Primordia were induced in the outer most cortical cell layers; (b) Determinate nodules formed; (c) Cluster of dividing cells at the apex of nodule; (d) High magnification of dividing cells at metaphase with intracellular bacteria inside. VB, vascular bundle; C, cluster of dividing cells; CN, central zone; S: symbiosome; Scale bars: a, 25µm; b-c, 75µm; d, 7.5µm.

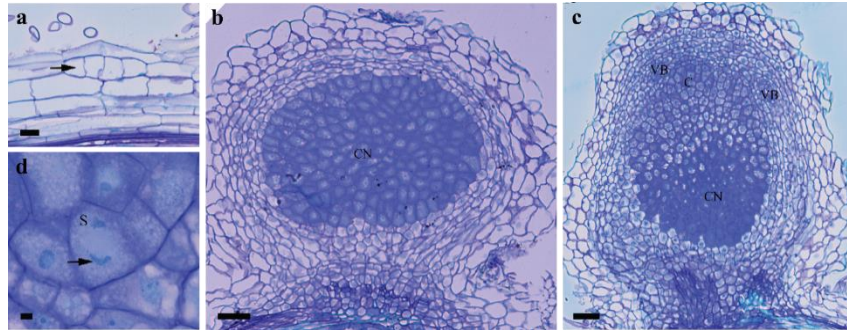


Fig. 5 Longitudinal section of primordia and nodules formed on *U. europaeus*. (a) Primordia were induced in the outer most cortical cell layers; (b) Determinate nodules formed; (c) Cluster of dividing cells at the apex of nodule; (d) High magnification of dividing cells at anaphase with intracellular bacteria inside. VB, vascular bundle; C, cluster of dividing cells; CN, central zone; S: symbiosome; Scale bars: a, 25 μ m; b-c, 75 μ m; d, 5 μ m.

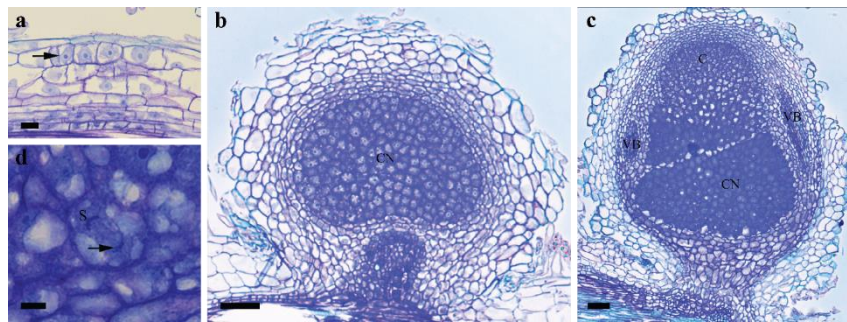


Fig. 6 Longitudinal section of primordia and nodules formed on *C. scoparius*. (a) Primordia were induced in the outer most cortical cell layers; (b) Determinate nodules formed; (c) Cluster of dividing cells at the apex of nodule; (d) High magnification of dividing cells at anaphase with intracellular bacteria inside. VB, vascular bundle; C, cluster of dividing cells; CN, central zone; S: symbiosome; Scale bars: a, 25 μ m; b-c, 75 μ m; d, 7.5 μ m.

apex of *Sophora* (Fig. 4d), *Ulex* (Fig. 5d), and *Cytisus* (Fig. 6d) nodules leading to an elongated nodule shape, which was similar to *Tephrosia* and *Indigofera* (Chapter 2). So the position of dividing cells can vary within this genistoids clade and the clusters of dividing cells at the basal-lateral side of nodules seems specific for *Lupinus* (Fig. 3c).

This suggests that the most recent common ancestor of the genistoids (node 22 in Fig. 2) as well as the most recent common ancestor of *Lupinus*, *Cytisus*, and *Ulex* (node 18 in Fig. 2) had a nodule developmental program that at early stages is similar to that of *Indigofera* and milletioids. Further, at later stages, secondary clusters of dividing infected cells at their apex and the aberrant basal-lateral position evolved within *Lupinus*. The lupinoid nodule is synapomorphic in genistoids and derived from the determinate nodule developmental program. So the development of genistoids nodules is different from the typical indeterminate nodule

development of the IRLC species but it is similar to that of the *Indigofera* and milletioids species. As the genistoids and dalbergioids have a common ancestor (node 23 in Fig. 2), it is interesting to compare nodule development of species within these 2 clades.

2.6 Nodule development in the dalbergioids

Within the dalbergioids nodule development has been best studied in *Aeschynomene* species (Bonaldi et al., 2011). The primordia are induced in the outermost cortical cell layer (e.g. *A. indica*) or in more inner located cortical cell layers (e.g. *A. afraspera*). Nodule development of *Aeschynomene* species (e.g. *A. indica*), in which the primordium is formed in the outermost cortical cell is very similar to that of *Indigofera*, milletioids, and genistoids clades. However, in general secondary clusters of dividing infected cells are not formed in these species (Bonaldi et al., 2011), like in milletioids species (e.g. soybean, cowpea). In *Aeschynomene* species (e.g. *A. afraspera*), in which the primordia are formed in more inner layers, primordia are formed with a typical determinate growth pattern and a non-infected apical meristem is not formed. However, in these species a new mechanism has evolved by which indeterminate growth can be established. In these species (e.g. *A. afraspera*) outer cortical cells are traversed by infection threads which grow towards the primordia that are formed in more inner located layers. An outgrowth tissue with giant cells containing infection threads develops in the outer cortical layers, but bacteria are not released from the infection threads (Bonaldi et al., 2011). From these dividing outer cortical cells new nodule lobes are formed. Whether these outgrowths are formed at the primordium stage or later stage and how long the infected cells remain mitotically active in these *Aeschynomene* nodules are unclear.

To study whether these two nodule developmental processes are common in the dalbergioids, I analysed one more species, *Stylosanthes guyanensis*, which represents a relative late branching lineage within the dalbergioids clade (Fig. 2). My analyses showed that primordia were induced in the outermost cortical cell layer at the basis of a lateral root, like in *A. indica* (Fig. 7a). A non-infected meristem is not formed at primordium stage (Fig. 7b). Infected primordium cells keep dividing (Fig. 7d) and the nodule primordium develops into a determinate nodule (Fig. 7c). This determinate nodule developmental process is similar to that of determinate nodule (Bergersen and Briggs, 1958).

The determinate aeschynomenoid nodule is considered to be synapomorphic for the dalbergioids. However, my studies showed that *S. guyanensis* form determinate nodule similar as in *Indigofera*, milletioids, and genistoids. Therefore, the aeschynomenoid nodule formed by *Aeschynomene* species is not synapomorphic for the dalbergioids. The nodule type with lobes formed in some *Aeschynomene* species (e.g. *A. afraspera*) is derived from the determinate nodule developmental program. The studied dalbergioids species form determinate nodule in a similar way as *Indigofera*, milletioids and genistoids species, and share a common ancestor (node 23 in Fig. 2).

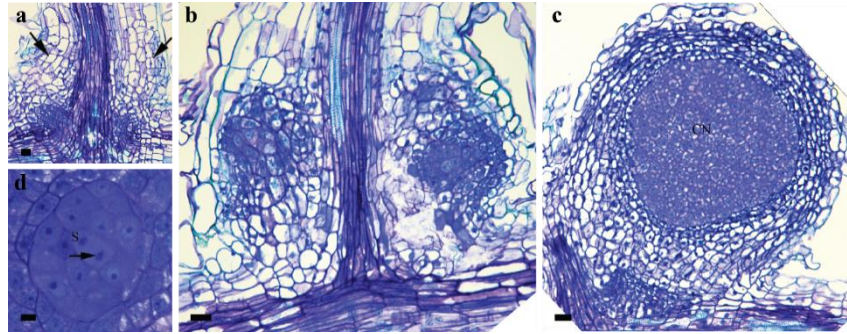


Fig. 7 Longitudinal section of primordia and nodules formed on *S. guyanensis*. (a) Primordia were induced in the outer most cortical cell layers at the basal of lateral root; (b) Non-infected meristem is not formed at primordium stage; (c) Determinate nodules formed; (d) High magnification of dividing cells at anaphase with intracellular bacteria inside. VB, vascular bundle; C, cluster of dividing cells; CN, central zone; S: symbiosome; Scale bars: a-c, 25µm; d, 7.5µm.

2.7 *Chamaecrista* nodule development

Based on the above described studies I conclude that the most recent common ancestor (node 23 in Fig. 2) of Papilionoideae had a “determinate” nodule development and most likely the ability to form secondary clusters of dividing infected cells, it formed most likely primordia by mitotic activation of the outermost cortical root cell layer. To support this conclusion, Tingting Xiao (unpublished) analysed the nodule developmental process of an ancient legume species *Chamaecrista fasciculata* from the Caesalpinioideae subfamily (Fig. 2).

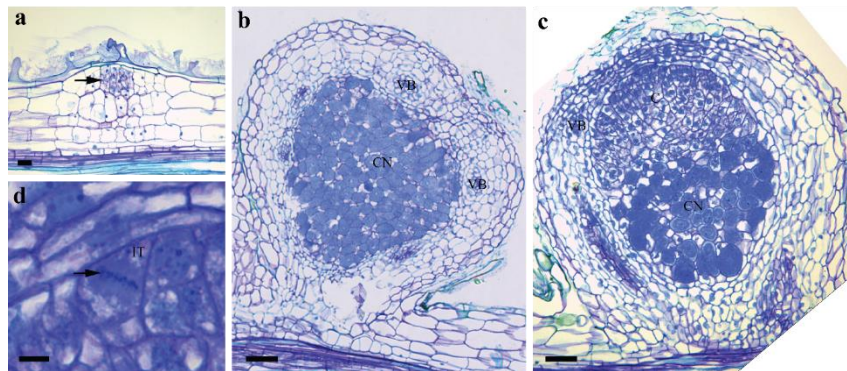


Fig. 8 Longitudinal section of primordia and nodules formed on *C. fasciculata*. (a) Primordia were induced in the outer most cortical cell layers; (b) Determinate nodules formed; (c) Cluster of dividing cells at the apex of nodule; (d) High magnification of dividing cells at metaphase with infection threads inside. VB, vascular bundle; C, cluster of dividing cells; CN, central zone; IT: infection thread; Scale bars: a, 25µm; b, 75µm; c, 50µm; d, 7.5µm. (Figures are from Tingting Xiao).

Nodule primordia of *Chamaecrista* (e.g. *C. fasciculata*) originate from the outermost cortical cells (Fig. 8a). Similarly, a meristem composed of non-infected cells is not formed at primordium stages and infected primordium cells maintain mitotically active to develop into a determinate nodule (Fig. 8b). At later stages, a group of small infected dividing cells is present at the apex (Fig. 8c, d). So *Chamaecrista* nodule development is very similar to that of species of the genistoids, *Indigofera*, milletioids and dalbergioids clades of the Papilionoideae subfamily. Therefore, I hypothesize that the common ancestor of the Papilionoideae forms determinate nodules that most likely have the ability to form secondary clusters of dividing cells. The persistent non-infected apical meristem formed in the IRLC clade might be unique for this clade. The sister clade of the IRLC is the robinoids. Therefore, it is of interest to determine whether some species within this clade can form nodules with such a persistent meristem.

2.8 Nodule development in the robinoids

Lotus is among the best studied robinoids species. In *Lotus*, nodule primordia are formed from the middle (third layer) cortex (Szczyglowski et al., 1998). However, cell division ceases at an early stage of development. Whether these primordium cells remain mitotically active when they become infected is unclear. The infection threads grow through the first and second cortical cell layer that do not divide. Therefore a non-infected apical meristem is not formed at the primordium stage. *Lotus* forms determinate nodules, but in contrast to most other species these are not formed from the outer most cortical cell layer. They are formed from the middle cortex like in some *Aeschynomene* species.

Sesbania shares with *Lotus* that nodule primordia are formed from the middle cortex. In contrast to *Lotus*, it can form nodules with either determinate or indeterminate growth behaviour (Ndoye et al., 1994; Fernández-lópez et al, 1998). The development of the determinate nodules is very similar to that of *Lotus* (and other species forming determinate nodules). Nodules with an indeterminate growth behaviour are formed when ethylene signalling is suppressed by a treatment with Ag_2SO_4 . This is a striking difference with *Medicago* that requires ethylene signaling for the formation of the meristem. The *Medicago Sickie* gene is an ortholog of *Arabidopsis Ein2*. A sickie mutant forms many small round nodules and these do not form an apical meristem (Penmetsa et al., 2008; Xiao et al., 2014). The indeterminate growing *Sesbania* nodules have a cluster of dividing cells at their apex (Ndoye et al., 1994; Fernández-lópez et al, 1998). However, whether it is a meristem as in the IRLC clade remains to be studied.

Another robinoids species which makes nodules with indeterminate growth behavior is *Robinia pseudoacacia*. I studied their development. The nodule primordia are induced in the outermost cortical cell layer and a non-infected meristem is not formed at the primordium stage (Fig. 9a, b). At a later stage of development, clusters of dividing infected cells with

infection threads are present at the apex leading to the indeterminate nodule growth (Fig. 9c, d).

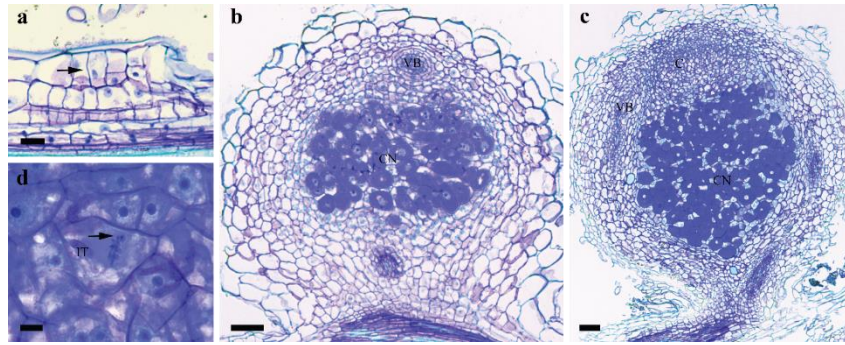


Fig. 9 Longitudinal section of primordia and nodules formed on *R. pseudoacacia*. (a) Primordia were induced in the outer most cortical cell layers; (b) Determinate nodules formed; (c) Cluster of dividing cells at the apex of nodule; (d) High magnification of dividing cells at metaphase with infection threads inside. VB, vascular bundle; C, cluster of dividing cells; CN, central zone; IT: infection thread; Scale bars: a, 25µm; b-c 75µm; d, 7.5µm.

A non-infected meristem is not formed in *Lotus* at the primordium stage nor in *Robinia*. Whether *Sesbania* forms an apical meristem during the primordium stage remains to be studied. If this is the case it most likely evolved independently from IRLC nodule meristem. So far a non-infected apical meristem has only been shown on IRLC species. I hypothesize that it is composed of stem cell like cells that can differentiate into different cell types. This property is shared with cells in other meristems, but cells in the secondary clusters of dividing cells do not share this property.

3. Evolution of NCR induced bacteroid development

NCR induced terminal differentiation has first been discovered in IRLC species like *Medicago* and *pea*. For a long time it was assumed that this trait was specific for this clade. However, the more recent discovery that *Aeschynomene* species also have NCR induced terminal differentiation showed that this was not correct. In both cases this is accompanied by increased bacteroid size and endoreduplication (Mergaert et al., 2003; Mergaert et al., 2006; Czernic et al., 2015). In Chapter 5, I showed that in *I. argentea* nodules, *B. elkanii* differentiates into enlarged bacteroids with a polyploid genome. This correlated with the expression of NCR genes. As the *Indigofera* bacteroids are also terminally differentiated, it is probable that this is induced by the NCR peptides. The *I. argentea* NCR peptides have different cysteine motifs as IRLC clade species or *Aeschynomene* species. This suggests they evolved independently from different defensin (Demina et al., 2013).

Enlarged bacteroids also occur in other clades within the Papilionoideae subfamily. For example, *Tephrosia* from millettoids and *Cytisus* from genistoids both contain enlarged and

polyploid bacteroids (Oono et al., 2010a), It is of interest to study whether these enlarged bacteroids are also triggered by NCR peptides. If this is the case the NCR induced bacteroid development evolved numerous times within the Papilionoideae subfamily and strongly suggest that this property is beneficial. It has been proposed that larger bacteroids improve the efficiency of N fixation (Oono et al., 2010b). Symbiosomes in general contain a single enlarged bacteroid whereas several bacteroids are present in a symbiosome when they do not enlarge. For this reason, enlarged bacteroids might be in better contact with the symbiosome membrane. This might increase nitrogen fixation efficiency. Another reason why the formation of enlarged bacteroids is beneficial is because it involves a terminal differentiation step by which the host can control the microsymbiont.

4. Evolution of indeterminate nodule and enlarged bacteroids in the Papilionoideae subfamily

Comparison of nodule development across the different clades of the Papilionoideae subfamily shows that the ancestral form is most likely the type we first observed in *Indigofera* and which also occurs in species of the genistoids and millettioids clades (Fig. 10). Its ancestral nature is underlined as the basal *Chamaecrista* from the Caesalpinioideae subfamily (Fig. 10) has a similar nodule development.

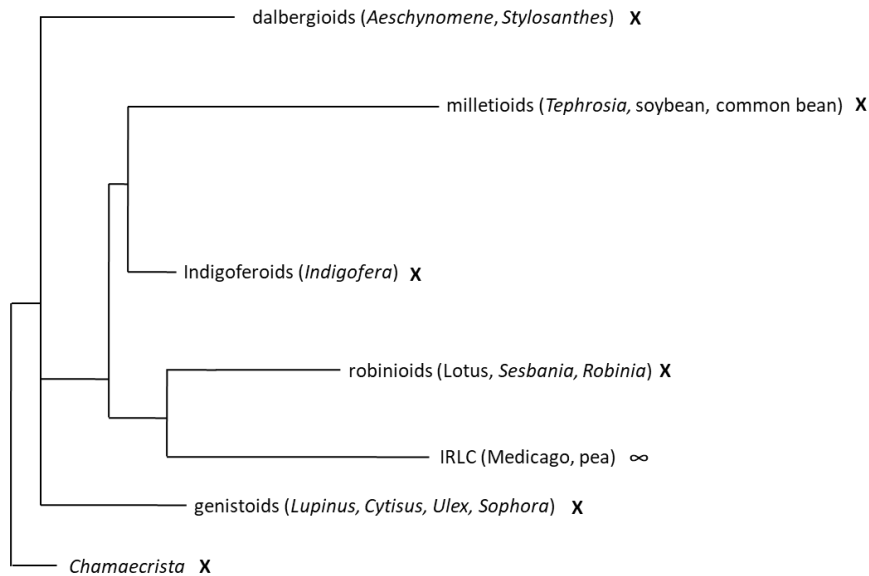


Fig.10 Species/genera belonging to the clades in the Papilionoideae subfamily are indicated (between brackets). Indeterminate (∞) or determinate (x) nodules formed by the species/genera in these clades is indicated next to the genus name.

The IRLC nodules are so far the only nodules with a genuine meristem. All other nodule types that occur within the Papilionoideae subfamily seem to be derived from the ancestral “determinate” nodule type. The determinate nodule type was first described for milletioids species (e.g. soybean). The determinate nodules that are formed by robinoids species have a very similar development as the determinate nodules of the milletioids. However, it remains to be studied whether *Sesbania* and *Robinia* species with nodules originate from the middle cortex nodules have a genuine meristem.

The formation of a secondary cluster of dividing cells provides the “determinate” nodules an indeterminate growth behavior. Within the Papilionoideae subfamily 2 additional mechanisms conferring indeterminate growth evolved. These are the mitotic activation of outer root cortical cells that were traversed by an infection thread in some *Aeschynomene* species (Fig. 11d) and the formation of an apical meristem in the IRLC species (Fig. 12b).

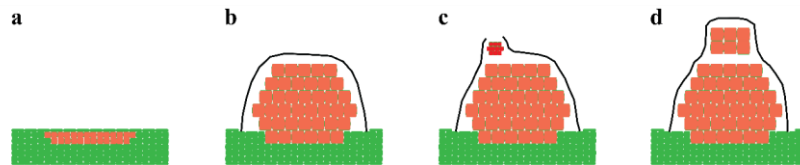


Fig. 11 Nodule development of ancestral nodule type. (a) Primordia are induced from outermost cortical layer; (b) The mitotic activity ceases at an early stage of development; (c) A secondary cluster of dividing infected cells is formed; (d) Ancestral nodule type II. Colors: red, outermost and middle cortical layers at primordium stage that differentiate into parts of central tissue; dark red, secondary cluster of dividing infected cells formed at later stage.

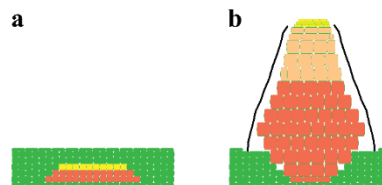


Fig. 12 Nodule development of indeterminate nodule type. (a) Primordia are induced from inner cortical layer; (b) non-infected meristem derived from the 3rd cortical layer. Color: yellow, 3rd cortical layer and future meristem; pink, infected central tissue derived from cell division of meristem; red, inner cortical layers and differentiate into a part of the central tissue.

The distribution of the different nodule types based on the results described in this thesis in the Papilionoideae subfamily is shown in Fig. 10. The IRLC clade that forms meristem containing (indeterminate) nodules evolved from an ancestor that forms “determinate” nodules that have the ability to establish secondary growth (node 28 in Fig. 10). So three mechanisms evolved within the Papilionoideae subfamily by which nodules can have indeterminate growth. The secondary cluster of dividing cells formed from differentiated

nodule cells, a genuine persistent meristem and the formation of a cluster of dividing cells containing an infection thread (Fig. 11c, d; Fig. 12b).

The independent evolution of indeterminate growth strategies suggests that this is beneficial for the symbiotic interaction. For example, to adapt to varying environmental conditions after drought, salt, or water stress; to create new infected cells when the older cells are senescent; or to support plant growth during different growth seasons of a perennial.

Our findings also imply that the general well accepted (textbook) names, indeterminate and determinate nodule types, have to be replaced by new ones. However, which names would be best will first require a discussion within the legume nodule community.

Materials and Methods:

Plant material and rhizobial strain isolation

This study included white lupin, *Sophora flavescens*, *Ulex europaeus*, *Cytisus scoparius*, *Stylosanthes guyanensis*, *Chamaecrista fasciculata*, and *Robinia pseudoacacia* (Table 1). Of which *U. europaeus*, *C. scoparius*, and *R. pseudoacacia* from Vreeken seeds, *S. flavescens* from Yinshan Jiao (China Agricultural University, China), *S. guyanensis* from Shu Chen (South China Agricultural University, China), *C. fasciculata* from Jeff J. Doyle (Cornell University, USA). All these species were grown in river sand and rhizobia were isolated and used as inoculants to induce nodule formation on the host plant as described in Chapter 3.

Table 1. Plant species and inoculated rhizobia strain.

Species	Bacteria strains
white lupin	<i>Bradyrhizobium japonicum</i> , L6;
<i>Sophora flavescens</i>	<i>Bradyrhizobium elkanii</i> , SA281
<i>Ulex europaeus</i>	<i>Bradyrhizobium elkanii</i> , SA281
<i>Cytisus scoparius</i>	<i>Bradyrhizobium elkanii</i> , SA281
<i>Stylosanthes guyanensis</i>	<i>Bradyrhizobium liaoningense</i> , Sg014
<i>Chamaecrista fasciculata</i>	Unpublished data
<i>Robinia pseudoacacia</i>	<i>Mesorhizobium</i> sp., Rob3

Root inoculation

Seed germination and plant inoculation were conducted as described in Chapter 2.

Tissue embedding, sectioning and section staining

Root segments and nodule sectioning were conducted as described in Chapter 4.

References

- Allen, O.N., and Allen, E.K. (1981). The Leguminosae: a source book of characteristics, uses and nodulation, London [etc.], GB: Macmillan [etc.].
- Bergersen, F.J., and Briggs, M.J. (1958). Studies on the bacterial component of soybean root nodules: Cytology and organization in the host tissue. *Journal of General Microbiology* 19(3): 482-490.
- Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., Goormachtig, S., and Giraud, E. (2011). Nodulation of *Aeschynomene afraspera* and *A. indica* by photosynthetic *Bradyrhizobium* sp. strain ORS285: The Nod-dependent versus the Nod-independent symbiotic interaction. *Molecular Plant-Microbe Interactions* 24(11): 1359-1371.
- Bryan, J.A., Berlyn, G.P., and Gordon, J.C. (1996). Toward a new concept of the evolution of symbiotic nitrogen fixation in the Leguminosae. *Plant and Soil* 186(1): 151-159.
- Callaham, D., Newcomb, W., Torrey, J.G., and Peterson, R.L. (1979). Root hair infection in actinomycete-induced root nodule initiation in *Casuarina*, *Myrica*, and *Comptonia*. *Botanical Gazette* 140: S1-S9.
- Czernic, P., Gully, D., Cartiaux, F., Moulin, L., Guefrachi, I., Patrel, D., Pierre, O., Fardoux, J., Chaintreuil, C., Nguyen, P., Gressent, F., Da Silva, C., Poulain, J., Wincker, P., Rofidal, V., Hem, S., Barriere, Q., Arrighi, J.F., Mergaert, P., and Giraud, E. (2015). Convergent evolution of endosymbiont differentiation in dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by Nodule-Specific Cysteine-Rich peptides. *Plant Physiology* 169(2): 1254-1265.
- Demina, I.V., Persson, T., Santos, P., Plaszczyca, M., and Pawlowski, K. (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): e72442.
- Fedorova, E.E., De Felipe, M.R., Pueyo, J.J., and Lucas, M.M. (2007). Conformation of cytoskeletal elements during the division of infected *Lupinus albus* L. nodule cells. *J Exp Bot* 58(8): 2225-2236.
- Fehér, D., and Bokor, R. (1926). Untersuchungen über die bakterielle Wurzelsymbiose einiger Leguminosenhölzer. *Planta* 2(4): 406-413.
- Fernández-lópez, M., Goormachtig, S., Gao, M., D'Haese, W., Van Montagu, M., and Holsters, M. (1998). Ethylene-mediated phenotypic plasticity in root nodule development on *Sesbania rostrata*. *Proc Natl Acad Sci U S A* 95(21): 12724-12728.
- Franssen, H.J., Xiao, T.T., Kulikova, O., Wan, X., Bisseling, T., Scheres, B., and Heidstra, R. (2015). Root developmental programs shape the *Medicago truncatula* nodule meristem. *Development* 142(17): 2941-2950.
- Friesner, G.M. (1926). Bacteria in the roots of *Gleditsia triacanthos* L. *Proc. Ind. Acad. Sci.* 34: 215-224.
- González-Sama, A., Lucas, M.M., De Felipe, M.R., and Pueyo, J.J. (2004). An unusual infection mechanism and nodule morphogenesis in white lupin (*Lupinus albus*). *New Phytologist* 163(2): 371-380.
- Gunawardena, U., and Hawes, M.C. (2002). Tissue specific localization of root infection by fungal pathogens: Role of root border cells. *Molecular Plant-Microbe Interactions* 15(11): 1128-1136.
- Gunawardena, U., Rodriguez, M., Straney, D., Romeo, J.T., VanEtten, H.D., and Hawes, M.C. (2005). Tissue-specific localization of pea root infection by *Nectria haematococca*. Mechanisms and consequences. *Plant Physiology* 137(4): 1363-1374.
- James, E.K., Sprent, J.I., Hay, G.T., and Minchin, F.R. (1993). The effect of irradiance on the recovery of soybean nodules from sodium chloride-induced senescence. *Journal of Experimental Botany* 44(263): 997-1005.
- James, E.K., Minchin, F.R., Iannetta, P.P.M., and Sprent J.I. (1997). Temporal relationships between nitrogenase and intercellular glycoprotein in developing white lupin nodules. *Annals of Botany* 79: 493-503.
- Lancelle, S.A., and Torrey, J.G. (1985). Early development of rhizobium-induced root-nodules of *Parasponia rigida*. II. Nodule morphogenesis and symbiotic development. *Canadian Journal of Botany-Revue Canadienne De Botanique* 63(1): 25-35.
- Lewis, G. (2005). Pp. 187-191. in: Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (eds.), *Legumes of the World*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- LPWG, Legume Phylogeny Working Group (2017). A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. *TAXON* 66 (1): 44-77.
- Marchetti, M., Catrice, O., Batut, J., and Masson-Boivin, C. (2011). *Cupriavidus taiwanensis* bacteroids in *Mimosa pudica* indeterminate nodules are not terminally differentiated. *Applied and Environmental Microbiology* 77(6): 2161-2164.
- Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., and Kondorosi, E. (2003). A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiology* 132(1): 161-173.

- Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., and Kondorosi, E. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* 103(13): 5230-5235.
- Meyerowitz, E.M. (1997). Genetic control of cell division patterns in developing plants. *Cell* 88(3): 299-308.
- Naisbitt, T., James, E.K., and Sprent, J.I. (1992). The evolutionary significance of the Legume genus *Chamaecrista*, as determined by nodule structure. *New Phytologist* 122(3): 487-492.
- Ndoye, I., De Billy, F., Vasse, J., Dreyfus, B., and Truchet, G. (1994). Root nodulation of *Sesbania rostrata*. *J Bacteriol* 176(4): 1060-1068.
- Oono, R., and Denison, R.F. (2010b). Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiology* 154(3): 1541-1548.
- Oono, R., Schmitt, I., Sprent, J.I., and Denison, R.F. (2010a). Multiple evolutionary origins of legume traits leading to extreme rhizobial differentiation. *New Phytologist* 187(2): 508-520.
- Pawlowski, K., and Bisseling, T. (1996). Rhizobial and actinorhizal symbioses: What are the shared features? *Plant Cell* 8(10): 1899-1913.
- Penmetsa, R.V., Uribe, P., Anderson, J., Lichtenzweig, J., Gish, J.C., Nam, Y.W., Engstrom, E., Xu, K., Sckisel, G., Pereira, M., Baek, J.M., Lopez-Meyer, M., Long, S.R., Harrison, M.J., Singh, K.B., Kiss, G.B., and Cook, D.R. (2008). The *Medicago truncatula* ortholog of Arabidopsis EIN2, *sickle*, is a negative regulator of symbiotic and pathogenic microbial associations. *Plant Journal* 55(4): 580-595.
- Price, G.D., Mohapatra, S.S., and Gresshoff, P.M. (1984). Structure of nodules formed by *Rhizobium* strain Anu289 in the nonlegume *Parasponia* and the legume siratro (*Macroptilium atropurpureum*). *Botanical Gazette* 145(4): 444-451.
- Roth L.E., and Stacey, G. (1989). Bacterium release into host cells of nitrogen-fixing soybean nodules: the symbiosome membrane comes from three sources. *European Journal of Cell Biology* 49(1): 13-23.
- Scheres, B., Wolkenfelt, H., Willemsen, V., Terlouw, M., Lawson, E., Dean, C., and Weisbeek, P. (1994). Embryonic origin of the *Arabidopsis* primary root and root meristem initials. *Development* 120(9): 2475-2487.
- Sprent, J.I. (2007). Evolving ideas of legume evolution and diversity: A taxonomic perspective on the occurrence of nodulation. *New Phytologist* 174(1): 11-25.
- Sprent, J.I., Ardley, J., and James, E.K., (2017). Biogeography of nodulated legumes and their nitrogen fixing symbionts. *New Phytologist* 215: 40-56.
- Szczygłowski, K., Shaw, R.S., Wopereis, J., Copeland, S., Hamburger, D., Kasiborski, B., Dazzo, F.B., and De Bruijn, F.J. (1998). Nodule organogenesis and symbiotic mutants of the model legume *Lotus japonicus*. *Molecular Plant-Microbe Interactions* 11(7): 684-697.
- Tang, C.X., Robson, A.D., Dilworth, M.J., and Kuo, J. (1992). Microscopic evidence on how iron deficiency limits nodule initiation in *Lupinus angustifolius* L. *New Phytologist* 121(3): 457-467.
- Van de Velde, W., Guerra, J.C.P., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., Mergaert, P., Kondorosi, E., Holsters, M., and Goormachtig, S. (2006). Aging in legume symbiosis. A molecular view on nodule senescence in *Medicago truncatula*. *Plant Physiology* 141(2): 711-720.
- Van Ghelue, M., Ribeiro, A., Solheim, B., Akkermans, A.D., Bisseling, T., and Pawlowski, K. (1996). Sucrose synthase and enolase expression in actinorhizal nodules of *Alnus glutinosa*: comparison with legume nodules. *Mol Gen Genet* 250(4): 437-446.
- 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(18): 3517-3528.

ACKNOWLEDGEMENTS

This thesis could never been successfully accomplished without the help from so many colleagues and friends. I would like to express gratitude and appreciation to everyone that has contributed or helped in any possible way during my PhD research.

First of all, thank you Ton, for giving me the opportunity to be a member of Molecular Biology. After five years I can still remember clearly how exciting I was when you emailed me the invitation letter. This leads me to a totally different life. Your optimistic leadership supported me to survive this PhD journey, and your wisdom make the struggles less hard with the way of thinking, organizing the experiments, presenting, writing and reading the literatures. I cannot image how this thesis could be finished without all your helps.

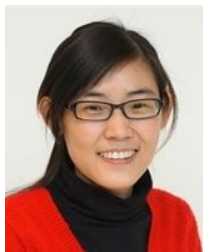
Thank you Henk, for leading me step by step in the lab at the very beginning of my PhD as a daily supervisor. You called me little girl at that time. I enjoy the time when you showed me how to do the PCR, how to work on the *Indigofera* transcriptome data and find out the ENOD2 gene, showed me the related papers to the project and supervise me on my first literature discussion. There are many more good times to be together with you.

Thank you Rene, you have tried to give me good suggestions and helps on the project, eventhough some of them I don't like so much. I cannot forget that before the first work discussion, I showed you the *Indigofera* nodule sections and felt hesidate to present the data, it is you strongly supported that I should do it. After that, Ton has interesting on the *Indigofera* nodule development, then the nodule evolution work started and end as a nice story.


Thank you Olga, for the ENOD2 *in situ* hybridization work on the *Indigofera* nodule and to check the *Indigofera* nodule primordium with the help of methylene blue. Robin, thank you for doing the *Indigofera* transcriptome data anlysis, this dataset helps a lot in the consecutive work as has been mentioned. Carolien, thank you or organizing and helping of the work in the lab, especially the *Indigofera* transformation work which has an excellent result better than I expected. Jan Verver, thank you for helping me on all the work about the DNA, RNA, PCR, cloning. I feel safty if you are there when I do the experiments as I always told you, "Jan Verver, I need you". Arjan, I would like to especially say, thank you so much, you are alwasy there whenever I need you, not only on the work but also the life. You helped me so much that I don't want to write down here but will remember forever. Rik, I enjoy the discussion with you on the experiments, music, and life. Especially I like the moments when you make fun of me, you are happy and also make me happy. You are one of my paranymph candidates, it's a pity that you cannot join.

And thanks to all the colleagues in the MOLBI lab and my lovely chinese friends.

CURRICULUM VITAE



Guiling Ren was born on the 13th of September 1986 in Xinxiang, Henan province, China. She finished her high school in 2005, and started her 4 years Bachelor education in Northwest A & F University in Yangling, China. During her study in the Horticulture, she explored the interesting towards the science. In 2009, she got the Bachelor degree and moved on to start her Master education in Zhejiang University in Hangzhou, China. There she worked on the Lilium breeding. In 2012, she obtained her Master degree and got the fellowship from China Scholarship Council (CSC) for supporting her PhD study in Wageningen University, the Netherlands. There she worked in the Laboratory of Molecular Biology under the supervision of Ton Bisseling and Rene Geurts. During her PhD study she mainly focused on the project of nodule evolution by analyzing the desert plant *Indigofera argentea*, and expanded to other related legume species.

<p style="text-align: center;">Education Statement of the Graduate School</p> <p style="text-align: center;">Experimental Plant Sciences</p>	
<p>Issued to: Gulling Ren</p> <p>Date: 23 January 2018</p> <p>Group: Laboratory of Molecular Biology</p> <p>University: Wageningen University & Research</p>	
	
1) Start-up phase	<i>date</i>
<ul style="list-style-type: none"> ▶ First presentation of your project <i>Title: Exploiting Natural Variation in the Heat and Desiccation Tolerance of the Indigofera-Rhizobium Symbiosis</i> 	04 Oct 2013
<ul style="list-style-type: none"> ▶ Writing or rewriting a project proposal <i>Title: Exploiting Natural Variation in the Heat and Desiccation Tolerance of the Indigofera-Rhizobium Symbiosis</i> 	10 Sep 2012
<ul style="list-style-type: none"> ▶ Writing a review or book chapter ▶ MSc courses ▶ Laboratory use of isotopes 	
<p style="text-align: right;"><i>Subtotal Start-up Phase</i> 7.5 credits*</p>	
2) Scientific Exposure	<i>date</i>
<ul style="list-style-type: none"> ▶ EPS PhD student days EPS PhD student day, Amsterdam, NL EPS PhD student day, Leiden, NL EPS PhD student day 'Get2Gether', Soest, NL European Plant Science Retreat for PhD students in Experimental Plant Sciences, Paris, France EPS PhD student day 'Get2Gether', Soest, NL 	30 Nov 2012 18 Nov 2013 29-30 Jan 2015 10-13 Jul 2015 28-29 Jan 2016
<ul style="list-style-type: none"> ▶ EPS Theme Symposia EPS Theme 2 Symposium "interactions between Plants and Biotic Agents, together with Willie Commelin Scholten Day, Utrecht, NL EPS Theme 4 Symposium 'Genome Biology', Wageningen, NL EPS Theme 4 Symposium 'Genome Biology', Wageningen, NL EPS Theme 3 Symposium 'Metabolism and Adaptation', Wageningen, NL EPS Theme 2 Symposium "interactions between Plants and Biotic Agents, together with Willie Commelin Scholten Day, Utrecht, NL EPS Theme 2 Symposium "interactions between Plants and Biotic Agents, together with Willie Commelin Scholten Day, Leiden, NL EPS Theme 1 Symposium 'Developmental Biology of Plants', Wageningen, NL 	24 Jan 2013 13 Dec 2013 03 Dec 2014 10 Feb 2015 20 Feb 2015 22 Jan 2016 21 Jan 2016
<ul style="list-style-type: none"> ▶ National meetings (e.g. Lunteren days) and other national platforms Annual meeting 'Experimental Plant Sciences', Lunteren, NL Annual meeting 'Experimental Plant Sciences', Lunteren, NL Annual meeting 'Experimental Plant Sciences', Lunteren, NL Annual meeting 'Experimental Plant Sciences', Lunteren, NL 	22-23 Apr 2013 14-15 Apr 2014 13-14 Apr 2015 11-12 Apr 2016
<ul style="list-style-type: none"> ▶ Seminars (series), workshops and symposia Plant Sciences Seminar: Ecological principles of plant-plant interactions: linking crops and natural systems Workshop 'Advanced Light Microscopy facilities at Wageningen UR' Plant Sciences Seminar: The use of Next generation sequencing in diagnostics and detection of plant pathogens EPS Flying Seminar Ortrun Mittelsten Scheid: 'Genetics and epigenetics: a complex relationship' Symposium 'Crop Pathology and Plant-Microbe Interactions' Flying Seminar Prof. dr. Jeff Doyle: 'Polyploidy in wild relatives of soybean and other legumes: systematics, comparative and functional genomics, and nodulation' EPS Flying Seminar Dr. Siobhan Brady: 'Regulation of root morphogenesis in tomato species in the face of a changing environment' WEES seminar: Julien Varaldi, Behaviour manipulation and horizontal gene transfer in a virus-parasitoid interaction EPS Flying Seminar dr. Rossana Henriques, 'And yet they oscillate: functional analysis of circadian long non-coding RNAs' WEES seminar Prof. Marc-André Selosse; 'Mycoheterotrophy and mixotrophy: plants eating mycorrhizal fungi' Seminar Dr. Jan Ruijter; 'Analysis of qPCR data. The use and usefulness of amplification curve analysis' 	11 Dec 2012 13 Jun 2013 11 Feb 2014 19 Nov 2014 08 May 2015 12 May 2015 09 Sep 2015 22 Oct 2015 16 Nov 2015 26 Nov 2015 14 Mar 2016
<ul style="list-style-type: none"> ▶ International symposia and congresses Rhizosphere 4, Maastricht, NL 12th European Nitrogen Fixation Conference, Budapest, Hungary 	21-25 Jun 2015 25-28 Aug 2016
<ul style="list-style-type: none"> ▶ Presentations <i>Pitch: Exploiting Natural Variation in Desiccation Tolerance of Rhizobia in Saudi Arabian Desert/Spring School 'Host-Microbe Interactomics'</i> <i>Poster: Exploiting Natural Variation in Desiccation Tolerance of Rhizobia in Saudi Arabian Desert/Spring School 'Host-Microbe Interactomics'</i> <i>Talk: The Evolution of determinate and indeterminate root nodule development in legumes/Theme 1</i> 	02-04 Jun 2014 02-04 Jun 2014 20 Jan 2016
<ul style="list-style-type: none"> ▶ IAB interview ▶ Excursions 	
<p style="text-align: right;"><i>Subtotal Scientific Exposure</i> 14.6 credits*</p>	

3) In-Depth Studies	
► EPS courses or other PhD courses	<u>date</u>
Course 'Bioinformatics - A User's Approach', Wageningen, NL	04-08 Mar 2013
Course 'Host-Microbe Interactomics', Wageningen, NL	02-04 Jun 2014
Course 'WIAS Advanced Statistics Course Design of Experiments', Wageningen, NL	07-09 Oct 2015
Course 'The Power of RNA seq', Wageningen, NL	10-12 Feb 2016
► Journal club	
participation in a literature discussion group at Molecular Biology	2012-2016
<i>Subtotal In-Depth Studies</i>	
<i>7.2 credits*</i>	
4) Personal development	
► Skill training courses	<u>date</u>
WGS course Scientific Writing, Wageningen, NL	Sep-Nov 2015
WGS course Project and Time Management, Wageningen, NL	May-Jun 2016
► Organisation of PhD students day, course or conference	
► Membership of Board, Committee or PhD council	
<i>Subtotal Personal Development</i>	
<i>3.3 credits*</i>	
TOTAL NUMBER OF CREDIT POINTS*	
32.6	
Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits	

* A credit represents a normative study load of 28 hours of study.

The research described in this thesis was performed at the Laboratory of Molecular Biology, Wageningen University, and financially supported by the China Scholarship Council (CSC) and Wageningen University.

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

Cover design: GVO drukkers & vormgevers B.V.

Layout: Guiling Ren

Printed by: GVO drukkers & vormgevers B.V., Ede, NL