Multiple molecular components contribute to genotype specific compatibility of the root nodule symbiosis

Dissertation

der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

vorgelegt von

Jasmin A. Gossmann

München

10. Januar 2011

Dekan: Prof. Dr. Benedikt Grothe

Erstgutachter: Prof. Dr. Martin Parniske

Zweitgutachter: Prof. Dr. Thomas Lahaye

Datum der Disputation: 17. März 2011

To my family and friends who have always believed in me.

Za moju obitelj i moje prijatelje koji su uvijek imali povjerenje u mene.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AM	Arbuscular mycorrhiza
At	*Arabidopsis thaliana*
AvrPto	Avirulence gene from *Pseudomonas syringae* pv. *tomato*
BAP	Benzlaminopurine
BAK1	BRI1-ASSOCIATED KINSASE
Bj	*Bradyrhizobium japonicum*
*Brady*NZP2309	*Bradyrhizobium* sp. strain NZP2309
CCAMK	CALCIUM CALMODULIN-DEPENDENT PROTEIN KINASE
CERK1	CHITIN ELICITOR RECEPTOR KINASE 1
CM	Co-cultivation medium
DNA	Deoxyribonucleic acid
DsRed	Discosoma sp. red fluorescent protein
DUF	Domain of unknown function
EDTA	Ethylenediaminetetraacetic acid
EIZ	Elongated infection threads
EMS	Ethyl methane-sulfonate
EOS	exooligosaccharides
EPS	Exopolysaccharides
F1	Filial generation 1
F2	Filial generation 2
FP	Fahraeus Plant medium
GFP	Green fluorescent protein
GUS	β-glucuronidase
IT	Infection thread
KCl	Potassium chloride
KD	Kinase domain
LB	Lysogeny broth
Lj	*Lotus japonicus*
LPS	Lipopolysaccharides

*LpLj*6.2	F1 generation hybrid 6.2 from the original cross *L. pedunculatus* LE306 x *Lj* Gifu
LysM	Lysine motif
MAMP	microbe-associated molecular pattern
*Meso*MAFF303099	*Mesorhizobium loti* strain MAFF303099
*Meso*DT3S	*Mesorhizobium loti* mutant strain DT3S
MP	Maximum parsimony
NAA	Naphthaleneacetic acid
NaCl	Sodium chloride
NF	Nod factor
NFR	NOD FACTOR RECEPTOR
NIN	NODULE INCEPTION
NJ	Neighbor-joining
nrITS	nuclear ribosomal internal transcribed spacer
OD	Optical density
PAMP	Pathogen-associated molecular patterns
PCR	Polymerase chain reaction
PTI	PAMP-triggered immunity
Pto	*Pseudomonas syringae* pv. *tomato*
RFP	Red fluorescent protein
RHC	Root hair curling
Rl-10.2.1	*Rhizobium leguminosarum* strain 10.2.1
Rlv	*Rhizobium leguminosarum* bv. *viciae*
RNS	Root nodule symbiosis
SDS	Sodium dodecyl sulfate
SP	Signal peptide
SSR	Simple sequence repeat
SYMRK	Symbiosis receptor-like kinase
T3SS	Type III secretion system
VIM	Naphthaleneacetic acid
w/v	weight/volume
X-Gal	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside
X-Gluc	5-bromo-4-chloro-3-indolyl-β-D-glucuronide

LIST OF UNITS

bp	Base pair
°C	Celsius
cm	Centimeter
h	Hour
kbp	Kilo base pair
L	Liter
m	Meter
min	Minutes
ml	Milliliter
mM	Millimolar

I SUMMARY

Two major factors conferring recognition specificity of the root nodule symbiosis between legumes and rhizobia are the bacterial perception systems for the plant derived flavonoid signals and the plant perception system for the rhizobial nod factors. To assess and characterize natural variation in the relevant genetic determinants in a model legume, we sampled endemic populations of *Lotus corniculatus* and *L. pedunculatus* and a range of rhizobial species from 21 sites across Europe. In a large nodulation assay we inoculated selected crossable diploid *Lotus* genotypes with 21 rhizobial strains, including newly collected isolates and reference strains. Using a new isolate of *Rhizobium leguminosarum*, we identified novel infection and organogenesis phenotypes and we uncovered variation in the symbiotic capability among diploid *Lotus* species and ecotypes that are normally obscured by optimally adapted *M. loti* strains. We found two distinct patterns of symbiotic GUS activity in the reporter line T90 that were correlated with different levels of bacterial infection compatibility. Using F1 hybrids from a cross between *L. pedunculatus* x *L. japonicus*, where both parental lines belong to contrasting infection and nodulation organogenesis compatibility groups, we found that plant-controlled selection and exclusion mechanisms are active in root nodule symbioses of *L. pedunculatus*. Inoculations with *M. loti* identified a negative regulatory mechanism controlling bacterial entry post infection thread formation and progression at the stage of bacterial uptake into host cells, dependent on the presence of the bacterial type III secretion system. By investigating the molecular evolution of an early infection specificity-encoding candidate gene, *NOD FACTOR RECEPTOR 5* (*NFR5*), from natural European *Lotus* populations, we found high levels of sequence diversity at the *NFR5* locus. We identified sites with signatures of positive selection in surface exposed loops of the LysM2 domain and the kinase domain. We also tested T90 expression in the background of symbiotic mutant lines and found that in hairy roots expressing a T265D gain-of-function construct of the CALCIUM CALMODULIN-DEPENDENT KINASE (CCAMK) the symbiotic GUS activity was induced in a rhizobium and nod factor independent manner and was absent in empty vector transformed roots of lines *ccamk-2*, *castor-2*, *symrk-10* and *nfr1-1*. The low level of non-symbiotic GUS activity that occurred in these mutants was detected in vacuum infiltrated seedlings. In summary these results indicate that rhizobium induced T90 expression is activated via the common *SYM* pathway but the non-symbiotic T90 expression is not.

ZUSAMMENFASSUNG

Zwei wesentlichen Faktoren, die der Wurzelknöllchensymbiose zwischen Leguminosen und Rhizobien Spezifität verleihen, sind einerseits die bakteriellen Perzeptionssysteme für pflanzenbürtige Flavonoidsignale und andererseits die pflanzlichen Perzeptionssysteme für die rhizobiellen Nod-Faktoren (NF). Um die natürliche Variation der relevanten genetischen Determinanten einer Modellleguminose zu untersuchen und zu charakterisieren, haben wir endemische *Lotus corniculatus* und *L. pedunculatus* Populationen sowie eine Reihe verschiedener Rhizobienarten an 21 Standorten quer durch Europa beprobt. Im Rahmen eines breit angelegten Nodulationsassays inokulierten wir ausgewählte diploide *Lotus*-Genotypen mit 21 Rhizobienstämmen einschließlich neu gesammelter Isolate und Referenzstämme. Mittels eines neuen Isolates von *Rhizobium leguminosarum* identifizierten wir neue Infektions- und Organogenesephänotypen, und wir entdeckten Variationen im symbiotischen Potential zwischen diploiden *Lotus*-Arten und Ökotypen, die optimal angepasste *M. loti* Stämme normalerweise verdecken. Wir fanden zwei distinkte symbiotische GUS-Aktivitätsmuster in der Reporterlinie T90. Diese korrelierten mit verschiedenen Stufen bakterieller Infektionskompatibilität. Mit Hilfe von F1-Hybriden einer Kreuzung zwischen *L. pedunculatus* x *L. japonicus*, wobei die parentalen Linien zu gegensätzlichen Infektions- bzw. Organogenese-Kompatibilitätsgruppen gehören, fanden wir dass pflanzengesteuerte Selektions- und Ausschlussmechanismen in Wurzelknöllchensymbiosen von *L. pedunculatus* aktiv sind. Anhand von Inokulationen mit *M. loti* identifizierten wir einen Negativ-Regulationsmechanismus, der, abhängig von der Gegenwart eines bakteriellen Typ III Sekretionssystems, den Eintritt von Bakterien nach der Infektionsschlauchausbildung im Stadium der Bakterienaufnahme in die Wirtszellen kontrolliert. Durch die Untersuchung der molekularen Evolution eines Kandidatengens für die Kodierung früher Infektionsspezifität, nämlich *NOD FAKTOR REZEPTOR 5* (*NFR5*) von natürlichen europäischen *Lotus*-Populationen, fanden wir hohe Sequenzdiversitätslevel im *NFR5* Lokus. Wir identifizierten Positionen mit Anzeichen positiver Selektion in oberflächenexponierten Schleifen der LysM2- sowie der Kinasedomäne. Wir testeten ebenfalls die T90-Expression im Hintergrund symbiotischer Mutantenlinien und fanden, dass die Exprimierung von T265D in „hairy roots", ein Funktionszunahme-Konstrukt der CALCIUM CALMODULIN-ABHÄNGIGEN KINASE (CCAMK), zur Induktion von symbioti-

scher GUS-Aktivität unabhängig von Rhizobien und NF führte. Sie war in Wurzeln der Linien *ccamk-2*, *castor-2*, *symrk-10* and *nfr1-1,* die mit dem leeren Vektor transformiert wurden, nicht vorhanden. Ein geringes Maß an nicht-symbiotischer GUS-Aktivität, die in Mutanten auftrat, wurde in vakuuminfiltrierten Keimlingen detektiert. Zusammenfassend weisen diese Ergebnisse darauf hin, dass die rhizobiuminduzierte T90-Expression im Gegensatz zur nicht-symbiotischen T90-Expression über den „common *SYM* pathway" induziert wird.

II   INTRODUCTION

II.1   Evolution and diversity of the root nodule symbiosis

Almost all flowering plant species that can form root nodule symbioses (RNS) with nitrogen fixing rhizobia belong to the Fabaceae family, which is one of the largest families in the plant kingdom. Fossil evidence and molecular clock analysis indicate that the Fabaceae most likely originated approximately 60 million years ago and the diversification of the major subfamilies - Caesalpinioideae, Mimosoideae and Papilionoideae (Lewis *et al.*, 2005) - took place in rapid succession, within about one to two million years following the origin of this family (Lavin *et al.*, 2005). Given the widespread ability within the family to form associations with nodulating bacteria, it is likely that this symbiotic ability dates back at least to the most recent common ancestor of this plant family.

The nodulating bacteria that form associations with legumes are α-proteobacteria and belong to the genera *Rhizobium, Mesorhizobium, Bradyrhizobium, Allorhizobium, Sinorrhizobium* and *Azorhizobium.* Recent studies of nodule isolates from mostly tropical legumes describe several β-Proteobacteria in the genera *Burkholderia* and *Cupriavidus* (syn. *Ralstonia*) that effectively nodulate some mimosoid and few papilionoid legumes in the genera *Mimosa, Cyclopia,* and *Rhynchosia* (Moulin *et al.*, 2001; Chen *et al.*, 2003; Chen *et al.*, 2005a; Chen *et al.*, 2005b; Elliott *et al.*, 2007; Garau *et al.*, 2009). The discovery of new genera of bacteria that engage in symbiotic associations with legumes reveals the large amount of variation found among nodulating rhizobia and challenges our previous view of legume-rhizobia coevolution. Partly, the diversity among nodulating bacteria may be due to the dynamic structure of bacterial genomes. For example, the organization of symbiosis genes within easily transmissible symbiosis islands or plasmids can permit the conversion of non-symbiotic bacteria into nitrogen-fixing plant endosymbionts in a single step (Sullivan *et al.*, 1995; Sullivan & Ronson, 1998; Marchetti *et al.*, 2010).

II.2   Host-rhizobium specificity

Nevertheless, the symbiotic interactions of legumes and rhizobia are highly specific. This is reflected in the different host ranges of rhizobia. Some rhizobia have narrow host-ranges; e.g. a cultivar-specific interaction was found in *Rhizobium leguminosarum* bv. *trifolii* and clover (Lewis-Henderson & Djordjevic, 1991). On the

other hand the strain *Rhizobium* sp. NGR234 has an exceptionally broad host-range and can nodulate several different legume species (e.g. *Glycine max, Phaseolus vulgaris, Robinia pseudoacacia*, and *Sesbania rostrata*) as well as non-legume *Parasponia andersonii (Webster et al., 1995; Broughton et al., 2000).*

Specificity is generally understood as infection compatibility of rhizobial strains with specific plant species. But, compatibility of rhizobia and plants encompasses all stages of the nodulation process including the fixation efficiency of rhizobia within the nodules (Ardourel et al., 1994; Banba et al., 2001; Okazaki et al., 2010). Thus, we understand 'incompatibility' as a phenotypic deviation from the nodulation phenotype of the natural host-rhizobium association or deduced and well-characterized combinations hereof. Incompatibility, defined in this way, can lead to impairment, deficiency, misperformance and other malfunctions or even blockage and termination of the RNS.

II.3   Developmental processes of nodulation

The RNS is the result of a complex succession of signaling events, which eventually lead to the formation of a root 'nodule'; a new organ in which the nitrogen-fixing bacteria are accommodated inside plant cells. The nodulation process can be divided into two developmental processes; infection and organogenesis. These two processes were successfully uncoupled using gain-of-function mutant alleles of the CALCIUM CALMODULIN-DEPENDENT KINASE (CCAMK) or the cytokinin receptor LHK1 (Madsen *et al.*, 2010). The same authors showed that host-encoded mechanisms control three alternative entry modes operating in the epidermis, the root cortex, and the single cell level. An earlier study on the semi aquatic legume *S. rostrata* found different rhizobial infection modes operating in a single legume, and those were correlated with different levels of structural requirements towards rhizobial nod factors (NF) (Goormachtig *et al.*, 2004). The three alternative host controlled entry processes are: 1) intercellular infection, 2) crack entry with or without infection threads (IT), and 3) root hair infection (Madsen *et al.*, 2010). The order of these three modes of infection may correspond to the timing of their evolutionary origin with bacterial intercellular infection without ITs as the earliest form of infection, progressing via a more derived mode of infection via cracks in the epidermis, to the most recently derived state involving root hair infection and transcellular ITs (Sprent,

2007; Madsen *et al.*, 2010). Local cell wall hydrolysis and invagination of the plasma membrane lead to formation of ITs. They contain bacteria and progress into the root cortex. In the primordium the IT branches, and the bacteria are released into the cytoplasm of plant cells, in which, surrounded by a peribacteroid membrane, they differentiate and reduce atmospheric dinitrogen to ammonium (Gage, 2004). Phenotypic analysis of plant and bacterial mutants indicate that compatibility between plant and microbial genotypes is required at several steps during rhizobial infection, allowing the plant to select for compatible rhizobia (Limpens *et al.*, 2003; Madsen *et al.*, 2003; Radutoiu *et al.*, 2003; Smit *et al.*, 2007).

Nodule organogenesis is characterized by bacterial and plant cell differentiation, leghemoglobin production and bacterial nitrogen fixation. There is evidence that the host determines the nodule morphology (Sprent, 2007; Sprent, 2008) and the differentiation of the bacteroids, but the effectiveness of N-fixation is largely controlled by the host-microsymbiont genotype combination (Mergaert *et al.*, 2006). However, Sachs *et al.* (2010) provide evidence for the ability of the legume host, *Lotus strigosus*, to constrain the infection and ensuing proliferation of a naturally occurring rhizobial cheater. At the point when the bacteria have already entered the host cells, the plant can only sanction cheating rhizobia with e.g. shortening the supply of nutrients or oxygen or early senescence (Kiers & Denison, 2008). One imperfection of the legume rhizobium symbiosis is the lack of a mechanistic connection between the extent of nitrogen fixation ability of a rhizobial strain and its compatibility. The latter being determined by bacterial NF structure and plant receptor recognition specificity.

II.4 Signal transduction

At least seven distinct mutants that are impaired in the symbiotic processes upstream of the calcium spiking were found using employing forward genetic tools including ethyl methane-sulfonate (EMS) mutagenesis or T-DNA insertions (Perry *et al.*, 2003; Miwa *et al.*, 2006; Perry *et al.*, 2009). Only two of these loci, namely *NOD FACTOR RECEPTOR1* (*NFR1*) and *NOD FACTOR RECEPTOR5* (*NFR5*), are RNS specific. *NFR1* and *NFR5* are required for calcium influx, membrane depolarization and transient alkalinization, as well as for the expression of nodulin genes (Madsen *et al.*,

2003; Radutoiu *et al.*, 2003; Miwa *et al.*, 2006). The *nfr1* and *nfr5* mutants show no root hair deformation upon *M. loti* inoculation (Radutoiu *et al.*, 2003).

Another six loci are involved in AM, as well as RNS; they are so-called common *SYM* genes (Kistner & Parniske, 2002). The common *SYM* mutants, namely *symrk*, *nup85*, *nup133*, *castor, pollux,* and the recently described *nena* mutant (Table 1), are impaired in the calcium spiking response and they show an increased root hair deformation phenotype compared to the wild-type upon incubation with rhizobia or NF (Bonfante *et al.*, 2000; Stracke *et al.*, 2002; Kanamori *et al.*, 2006; Miwa *et al.*, 2006; Groth *et al.*, 2010). These mutants fail to curl at the root hair tip, although this has not been described for the *nena* mutants. These early impairments suggest that the NF signaling pathway is blocked downstream of NF perception and that the calcium spiking signal is crucial for the root hair curling (RHC), followed by bacterial entrapment.

Table 1: Common *SYM* genes (including *NENA*) required for AM and RNS.

Gene/ Allele	Nod/ AM phenotype	Reference
SYMRK (*SYM2*)[a]	RHC$^-$, spiking$^-$	(Endre *et al.*, 2002; Stracke *et al.*, 2002)
NUP85 (*SYM24, SYM73* and *SYM85*)[a]	RHC$^-$, spiking$^-$	(Saito *et al.*, 2007)
NUP133 (*SYM3* and *SYM45*)[a]	RHC$^-$, spiking$^-$	(Kanamori *et al.*, 2006)
NENA[b]	RHC$^-$, spiking$^{(-)}$	(Groth *et al.*, 2010)
CASTOR (*SYM4* and *SYM71*)[a]	RHC$^-$, spiking$^-$	(Bonfante *et al.*, 2000; Ovtsyna *et al.*, 2005)
POLLUX (SYM23 and SYM86)[a]		(Kistner *et al.*, 2005)
CCAMK (SYM15 and SYM72)[a]	RHC$^-$, spiking$^+$	(Lévy *et al.*, 2004; Mitra *et al.*, 2004; Ovtsyna *et al.*, 2005; Tirichine *et al.*, 2006)
CYCLOPS (*SYM82*, *SYM6*, and *SYM30*)[a]	IT$^-$, spiking$^+$	(Messinese *et al.*, 2007; Yano *et al.*, 2008)

[a] previous name; [b] not described in the originally defined common *SYM* genes but shows the same early responses; RHC$^-$, no root hair curling; IT$^-$, blockage of infection thread initiation; spiking$^-$, no Ca^{2+} spiking; spiking$^+$, Ca^{2+} spiking.

A recent investigation with spontaneous nodulation mutants of *L. japonicus* and 16 genes involved in the RNS supports a model in which NFR1 and NFR5 signaling leads to RHC and IT formation, while the SYMBIOSIS RECEPTOR-LIKE KINASE (SYMRK), together with the cation channels CASTOR and POLLUX (Ané *et al.*, 2004; Imaizumi-Anraku *et al.*, 2005; Charpentier *et al.*, 2008), and the

nucleoporins NUP85 and NUP133 (Kanamori et al., 2006; Saito et al., 2007) function upstream of nuclear calcium spiking and are required for the generation of the spiking (Madsen *et al.*, 2010).

The *SYMRK* gene of *L. japonicus* encodes a transmembrane receptor-like kinase with an ectodomain comprising three leucine-rich repeat motifs and an extended N-terminal domain of unknown function (DUF) (Endre *et al.*, 2002; Stracke *et al.*, 2002). *Lj-SYMRK* can fully restore RNS of *Medicago* with *S. meliloti*, and *Os-SYMRK*, which lacks almost the entire DUF domain and one leucine repeat, successfully complements *Lj-symrk-10* mutants (Markmann *et al.*, 2008). This demonstrates that *SYMRK* is neither involved in determining legume–rhizobium specificity, nor in mediating specific developmental processes during nodule organogenesis.

Two components act downstream of the calcium spiking, CCAMK (Lévy *et al.*, 2004; Mitra *et al.*, 2004; Tirichine *et al.*, 2006) and a nuclear localized coil–coil protein CYCLOPS (Yano *et al.*, 2008), which reads out the spiking signature leading to organogenesis (Lévy *et al.*, 2004; Mitra *et al.*, 2004; Tirichine *et al.*, 2006; Madsen *et al.*, 2010). Mutant of *cyclops* transformed with the previously described T265D CCAMK gain-of-function construct form nodules devoid of ITs (Yano *et al.*, 2008; Madsen *et al.*, 2010). This demonstrates that CYCLOPS is required for CCAMK cross-signaling from the organogenesis pathway to the infection pathway (Madsen *et al.*, 2010).

II.5   NF signaling

NFs act at nano- to picomolar concentrations and serve as a specificity signal in RNS. They induce several responses in root hair cells, including intra- and extracellular alkalinization, membrane potential depolarization, changes in ion fluxes, and early nodulin gene expression. NFs are also essential for the deformation of root hairs and for the formation of ITs. High concentrations of NFs can elicit a calcium influx (Miwa *et al.*, 2006). However, physiological active concentrations are probably reached, when the bacteria are entrapped prior to the IT formation. A study of IT initiation in vetch by nodulation mutants of *Rhizobium leguminosarum* identifies the calcium influx as a necessary signal for the bacterial entry mode via the ITs (Walker & Downie, 2000). In the same study in a rhizobial mutant producing the simplest NFs

with a $C_{18:1}$ acyl chain, the expression of the *R. leguminosarum* gene *nodO* compensates for the lack of appropriate NF production and induces IT formation. The proposed role of NodO of *R. leguminosarum*, a secreted protein that can form cation-selective channels in lipid bilayers, is the stimulation of an ion flux such as the calcium influx across the host plasma membrane which could be sufficient for the induction of IT formations (Sutton *et al.*, 1994; Walker & Downie, 2000). However, the induced ITs are irregular in shape and arrest prematurely. The authors suggest that the initial entry, but not IT progression, can occur with minimal structural requirements of the NF, thus, with low host specificity. An earlier study of nodulation phenotypes induced in *Medicago* species by *S. meliloti* mutants, *nodF, nodE* and *nodL,* demonstrates that the structural requirements of NFs are higher for infection site formation, for IT formation and for bacterial entry than they are for the elicitation of developmental plant responses such as cell wall tip growth in trichoblasts (curling or swelling) and other epidermal cells as well as cortical cell activation (Ardourel *et al.*, 1994; Smit *et al.*, 2007). Ardourel *et al.* (1994) therefore suggested a two receptor model, comprising at least one signaling receptor with less stringent requirements towards NF structure, controlling developmental responses, and an entry receptor with higher NF structural requirements, controlling infection site and IT formation as well as bacterial entry. Both studies identify specific steps of the infection process with differing requirements towards rhizobial NF structure, and suggest that the root hair infection process requires progressively demanding NF structures. These observations strengthen the hypothesis that a structurally specific NF is necessary to pass a later "checkpoint" during infection – the proposed entry receptor or a third receptor that controls the IT progression. In the first model two receptor complexes, named signaling and entry receptor complexes, are required for rhizobial entry into the host cells. In the second model at least three receptor complexes (perception, entry and progression receptor complexes) are required for rhizobial infection.

II.6   NF signaling a key determinant of symbiotic specificity

The nodulation process requires reciprocal communication between the legume host and rhizobia through the exchange of molecular signals by both partners. If the partners are compatible, the signal cross-talk eventually leads to the formation of a new organ: the root nodule, which harbors the nitrogen-fixing rhizobia.

Rhizobial exo- and lipopolysaccharides (EPS and LPS), cyclic glucans, and effector proteins have demonstrable importance in facilitating the infection process (Perret *et al.*, 2000; Broughton *et al.*, 2003; Den Herder & Parniske, 2009; Kambara *et al.*, 2009; Okazaki *et al.*, 2010). Flavonoids are plant-derived phenolic signaling molecules. Different legumes release different mixtures of flavonoids into the rhizosphere. Upon flavonoid perception, rhizobial *nod*-genes are induced and lead to the production and secretion of NFs and effector proteins (Zehner *et al.*, 2008). However, the key determinants of the high degree of host specificity are the bacterial perceptions systems for plant derived flavonoids and bacterial produced NFs.

In *L. japonicus* two NOD FACTOR RECEPTORS, NFR1 and NFR5, are involved in early stages of the nodulation process for NF perception and activation of a signal cascade, and thus control bacterial infection (Madsen *et al.*, 2003; Radutoiu *et al.*, 2003). The *M. truncatula* NOD FACTOR PERCEPTION (NFP) protein, the ortholog of *Lj*-NFR5, was suggested to act as NF signaling receptor, and an additional LysM DOMAIN-CONTAINING RECEPTOR-LIKE KINASE (LYK3) probably functions as entry receptor by specifically controlling the bacterial entry via ITs in a NF dependent manner (Limpens *et al.*, 2003; Smit *et al.*, 2007). In *Lotus* the existence of an entry receptor controlling NF structure-specific bacterial infection via ITs is generally assumed (Ardourel *et al.*, 1994). However, until now no entry rector has been identified in *L. japonicus*. The two candidate genes, *NFR1* and *NFR5* encode trans-membrane receptor-like kinases (RLKs) that contain three extracellular LysM domains (Gough, 2003; Madsen *et al.*, 2003; Radutoiu *et al.*, 2003; Oldroyd & Downie, 2004). The transfer of *Lj-NFR1* and *Lj-NFR5* to *M. truncatula* enabled the transformants to recognize the *L. japonicus* symbiont *M. loti*, in addition to the *M. truncatula* symbiont *S. meliloti* (Radutoiu *et al.*, 2007).

A single mutation in the LysM2 domain of *NFR5* leading to an amino acid substitution (L188A) affected the nodulation efficiency of a *Lj-nfr5* mutant in combination with the structurally altered NFs produced by the genetically engineered rhizobium strain DZL (Radutoiu et al., 2007). These results suggest a direct binding of the NF to the LysM2 domain. However, direct binding to NFR1 or NFR5 has not yet been demonstrated. In-vitro binding was demonstrated for chitin and chitin oligosaccharides to LysM domain containing membrane proteins, *Arabidopsis thaliana* CHITIN ELICITOR RECEPTOR KINASE 1 (CERK1) and the kinase lacking rice CHITIN ELICITOR-BINDING PROTEIN (CEBiP), respectively (Kaku

et al., 2006; Iizasa *et al.*, 2010). The *Lj-NFR1* ortholog *At-CERK1* has been shown to be essential for chitin signaling, indicating an evolutionary relationship between chitin and NF perception (Zhu *et al.*, 2006; Miya *et al.*, 2007; Wan *et al.*, 2008). This hypothesis is also strengthened by the structural resemblance of chitin oligomers and NFs as lipo-chitooligosaccharides (LCO). Recently it has been shown that the only representative of the *NFR5* gene family in the non-legume *Parasponia* is not only required for bacterial entry but also for AM symbiosis (Camp *et al.*, 2010). Additionally, the Myc-factor structure has recently been identified as a lipochitooligosaccharide signal (Maillet *et al.*, 2011). These findings lend support to the long-standing hypothesis that AM fungi produce chitin-related signaling molecules to activate the common *SYM* pathway. Consequently, rhizobia probably acquired the ability to "hitchhike" on the ancient intracellular uptake program of AM by evolving signaling molecules that trigger the same or closely related symbiosis receptors (Markmann & Parniske, 2009).

II.7   Legume hosts choose rhizobial symbionts based on NF structure

Major factors conferring infection compatibility (specificity) to the RNS are the bacterial perception systems for the plant derived flavonoid signals and the rhizobial NFs. Incompatibilities of both symbiosis partner at the level of flavonoid or NF perception lead to a Nod⁻ phenotype (Horvath *et al.*, 1987; Spaink, H *et al.*, 1987; Long, 1996; López-Lara *et al.*, 1996). Rhizobia that produce incompatible NF structures are not selected by the legume host.

Legumes have evolved a wide range of nodule structures (e.g. determinate or indeterminate nodules, and root or stem nodules). The different nodule structures are congruent with the taxonomy of legumes, indicating that the legume host determines the nodule morphology (Sprent, 2007; Sprent, 2008). While legume hosts also determine the differentiation of the bacteroids, the effectiveness of N-fixation is largely controlled by the host-microsymbiont genotype combination (Mergaert *et al.*, 2006). For example, a N-fixation mutant (*nif*H) of the symbiont *M. loti* is able to induce infected, but non-fixing nodules on its corresponding host, *L. japonicus* (Ooki *et al.*, 2005). Additional legume-rhizobium genotype combinations display later incompatibilities at stages post organogenesis and infection. For example, *R. etli* induces nodules on *L. japonicus* that senesce prematurely (Banba *et al.*, 2001). The

formation of inefficient non-fixing nodules is a costly investment for the plant. Sachs *et al.* (2010) provide evidence for the ability of the legume host, *Lotus strigosus*, to constrain the infection and later proliferation of a naturally occurring rhizobial cheater. The mode of action of the host control of RNS is still obscure and might involve feedback loops and additional bacterial molecules. Using a type III secretion (T3SS) mutant of *M. loti* strain MAFF303099, it was shown that the T3SS positively affects nodulation on *L. corniculatus* subsp. *frondosus* and *L. filicaulis*, but negatively affects nodulation on *L. halophilus* (Okazaki *et al.*, 2010). Similarly, exooligosaccharides (EOS) of *Rhizobium* sp. NGR234 have no effect on the nodulation on *Vicia unguiculata*, but a negative effect on nodulation on *Leucaena leucocephala* (Staehelin *et al.*, 2006). However, the genetic basis of these incompatibilities at nodulation stages post organogenesis is only poorly understood. It is not yet known how legume plants perceive compatibility factors other than NFs.

The dilemma of the species-specific recognition of rhizobia by the legume host is that the plant selects the bacterial partner before the bacteria start fixing nitrogen. Den Herder & Parniske (2009) discuss implications of this imperfect selection, which provides an opportunity for inefficient, parasitic bacteria to induce nodulation and enter the plant host.

II.8   Selective forces that drive NF diversification and RNS specificity

Not all symbiotic interactions are mutualistic; some rhizobia are cheaters and consume carbohydrates from the plant and provide no or only low amounts of fixed nitrogen in exchange. Within a single legume the benefit to the plant from the rhizobia association can vary more than 10-fold (Burdon et al., 1999). The difference in nitrogen fixing rates by bacteria is important not only to the plant, but also for truly mutualistic (nitrogen-fixing) bacteria, which are in competition with non-fixing strains, so called "cheaters". Competition among strains for plant colonization provides a source of natural selection on both partners to establish a certain level of exclusiveness (Kiers & Denison, 2008). The ability of the host to distinguish between beneficial bacteria that are able to fix nitrogen versus parasitic bacteria that lack the ability to fix nitrogen is essential. But the plant cannot always distinguish between truly mutualistic rhizobia and bacterial cheaters, especially because the plant chooses the rhizobia before these actually start fixing nitrogen. The competitiveness of a non-

nitrogen fixing *Bradyrhizobium japonicum* nitrogenase mutant with its wild-type was demonstrated for nodulation of soybean nodules (Hahn & Studer, 1986; Parniske *et al.*, 1991). Over the long term, the formation of root nodules harboring low fixing rhizobia are costly to the legume hosts and could drive the evolution of exclusion of many potential symbionts. In fact, some legume lineages have independently lost the ability to nodulate (Sprent, 2007), maybe because the cost-benefit ratio is too high when nodulation does not contribute to the plant fitness (Foster & Wenseleers, 2006). By mimicking nitrogen-fixing rhizobia, bacterial cheaters force the plants to be selective towards non-cheaters and probably drive the evolution and diversification of specificity signals and the corresponding receptors recognizing these signals. Presuming binding of NFs to the LysM domains of NFRs and taking into account, that the NF structure is a specificity determinant for the bacterial infection, the LysM domains are potential targets of selective forces and have probably coevolved with bacterial nod-genes. Given the long time ever since nodulation has evolved, it can be assumed that a selective force acting on signals constituting the symbiotic communication and their encoding genes has lead to a coevolution of legumes and rhizobia. Indeed, there is evidence for coevolution in some legume-rhizobium associations, as could be demonstrated by cross-inoculation experiments with three regionally differing *Rhizobium etli* strains and their preferential nodulation of regional *P. vulgaris* (Aguilar et al., 2004).

Intriguingly, there is less evidence for specificity in the arbuscular mycorrhiza (AM) endosymbiosis (Hartmann *et al.*, 2009), in spite of the similarity of the signals used by both microorganisms, the bacterial NF signal and the fungal Myc-factor signal (Maillet *et al.*, 2011). In a recent publication, Camp *et al.* (2010) provide evidence that the NF perception mechanism is recruited from the widespread more ancient AM symbiosis as a result of a gene duplication event. Both receptors activate the common SYM pathway that is required in both symbioses (Parniske, 2008). A possible explanation for the different levels of specificity of both root symbioses could be the great potential of rhizobial genomes for rapid adaptation to new host species. The rhizobium evolution can proceed rapidly, because of a high level of plasticity, which characterizes their genomes. In contrast to the more conserved core genes in their chromosome, they carry accessory genes that allow them to flexibly adapt to different ecological niches. Accordingly, the *nod*-genes and the *fix*-genes are in close proximity on *sym*-plasmids (*Rhizobium leguminosarum*) or in symbiosis

islands (*B. japonicum* and *Mesorhizobium loti*). The horizontal transfer of these genes was found between indigenous *Sinorhizobium fredii* and introduced *B. japonicum* (Barcellos *et al.*, 2007). The possibility of horizontally transferred nodulation genes between non-cheaters and cheaters puts extra pressure onto the host to be discriminative between rhizobia and calls for the evolution of compatibility check-points at different stages during the infection as well as the fixation process.

II.9   European *Lotus* species form distinct RNS compatibility groups

The genus *Lotus* comprises about 130 species native to Europe, Asia, Africa, Australia and some islands (Degtjareva *et al.*, 2008). Its members are adapted to a wide range of habitats, from coastal environments and wetlands to high altitudes.

The two *Lotus* species, *L. pedunculatus* and *L. corniculatus* are widely distributed throughout Europe (Fig. 1). A nrITS sequence based phylogenetic work from Degtjareva et al. (2008) shows that the diploid and tetraploid members of the *Lotus corniculatus* complex, namely the species *L. burttii*, *L. filicaulis*, *L. japonicus*, *L. glaber*, and *L. corniculatus* are closely related to each other. *L. pedunculatus* together with some other species forms a sister clade to the *L. corniculatus* clade, indicating that *L. pedunculatus* and the *L. corniculatus* complex share a recent common ancestor. Diploid and self-fertile species of the *L. corniculatus* complex, *L. burttii*, *L. filicaulis*, and *L. japonicus* can be crossed to each other, which is important with respect to the employment of classical genetic methods (Somaroo & Grant, 1971). *L. japonicus*, along with *M. truncatula*, serve as model organisms for legume plants and root symbioses. *L. japonicus* can form two root symbioses, RNS and AM symbiosis. This and several features including self-fertility, diploidy, a relatively small genome size, a short generation time, make it a valuable species for classical and molecular genetics (Handberg & Stougaard, 1992).

Of the 130 *Lotus* species worldwide four species have been domesticated and improved through selection and plant breeding*: Lotus corniculatus, L. pedunculatus, L. glaber* and *L. subbiflorus (Díaz, P et al., 2005; Degtjareva et al., 2008). L. japonicus* is closely related to the agriculturally relevant species *L. corniculatus,* but more distantly related to *Lotus pedunculatus* and *Lotus subbiflorus (Swanson et al., 1990; Degtjareva et al., 2008).*

Nodulation specificity in general, is consistently more stringent in temperate legumes than in tropical legumes (Sprent, 2007). For example, within the genus *Lotus* there are two known compatibility groups: *i. B. japonicum* effectively nodulates *L. subbiflorus* and *L. pedunculatus*, and *ii. M. loti* effectively nodulates *L. tenuis*, *L. corniculatus*, and *L. japonicus* (Irisarri *et al.*, 1996; Saeki & Kouchi, 2000). This makes *Lotus* an excellent model to investigate specificity and its genetic basis.

In cross inoculation experiments, *Bradyrhizobium* sp. induces infected, though non-fixing nodules on *L. japonicus*, while *L. pedunculatus* blocks the cross-inoculated *M. loti* already at the stage of infection (Bek *et al.*, 2010). These two stages of compatibility, the early infection specificity and the late fixation-efficiency, make *Lotus* an excellent model to investigate the molecular basis of nodulation incompatibility, which is still obscure. Swapping experiments with *Lj-NFR1* and *Lj-NFR5* chimeric receptors containing the ectodomain either of *L. pedunculatus* or *L. japonicus* failed to identify a correlation of NF structure and NFR ectodomains with the infection blockage of *M. loti* by *L. pedunculatus* (Bek et al., 2010). Thus, the nature of the infection blockage of *M. loti* by *L. pedunculatus* remains unknown, as does the contribution of NFR5, the potential signaling receptor.

II.10 T90; a transgenic *L. japonicus* Gifu GUS reporter line for AM and RNS

II.10.1 Symbiotic GUS activity of reporter line T90

In the frame of a promoter tagging experiment the plant line T90 was generated (Webb *et al.*, 2000). This line carries the promoterless reporter gene *uidA* (β-glucuronidase; GUS) from *E. coli*, a T-DNA insertion 1.3 kb upstream of the gene *CBP1* (calcium-binding-protein). The T90 line expresses GUS specifically in roots and/or nodules and is inducible upon rhizobia and AM fungi treatment (Webb *et al.*, 2000; Kistner *et al.*, 2005). However, the *CBP1* gene is probably not essential for the establishment of RNS. The symbiotic GUS activity is induced in the rhizodermis and the root hairs. During the RNS the GUS activity is progressively restricted to the developing nodules and vanishes from the rest of the root. Two wild-type strains of *M. loti*, namely NZP2235 and NZP2037, induce the same GUS activity pattern on the reporter line T90, while PN4047, a non-nodulating NodC mutant derived from NZP2037, which is unable to produce NFs, failed to induce GUS activity (Chua *et al.*,

1985; Scott *et al.*, 1996; Webb *et al.*, 2000). This result highlights NFs as elicitors of the symbiotic GUS induction. Intriguingly, the T90 GUS expression is also induced upon inoculation with AM fungus *G. intraradices* (Kistner *et al.*, 2005). Although not known, it can be speculated that T90 is also Myc-factor inducible. The earliest time points showing GUS reporter activity in the T90 line tested are 24h upon *M. loti* inoculation and one week upon co-cultivation with *G. intraradices.* F3 progeny of *castor-2* x T90 and *ccamk-2* x T90 were generated by Kistner (2005) and showed no GUS staining after inoculation with *M. loti* or *G. intraradices*.

II.10.2 T90 as quick readout for root epidermis transfection assays

It is a lengthy process to generate stably transformed transgenic plants. The selection of homozygous lines can take several months. In contrast, the hairy root transformation is much faster, but still takes several weeks. The root hair transformation, a well-established technique in the model organisms, *L. japonicus* and *M. truncatula*, is used to investigate AM and the RNS (Stiller *et al.*, 1997; Martiran *et al.*, 1999; Boisson-Dernier *et al.*, 2001; Díaz, C *et al.*, 2005). However, a drawback of this transformation is the fairly demanding technical handling of the seedlings during the procedure, which leads to a restriction in numbers of constructs and plants that can be transformed at a time.

Considerable progress has been achieved in transformation of the model plant *A. thaliana.* Among these are the *A. tumefaciens*-mediated vacuum infiltration method (Bechtold *et al.*, 1993) and the co-cultivation method using *A. rhizogenes* (Campanoni *et al.*, 2007). The latter describes the efficient visualization of fluorescent proteins in *A. thaliana* root epidermis. Compared to the hairy root transformation technique, which is demanding and very time-consuming, the transient vacuum infiltration technique is quick and straightforward. A well-functioning vacuum infiltration system would allow large assays to test a wide range of constructs and substances for their symbiotic signaling capacity. For example different NFR1 and NFR5 alleles or orthologs can be quickly tested for their symbiont specific effects, but many other applications are conceivable. The vacuum infiltration method coupled with a quick read-out method, e.g. the T90 GUS staining, can enable a wider range of experimental assays. The reporter line T90 is a good candidate for the combination with a transient epidermal transfection system, because it can be used to test for both root symbioses.

25

II.11 Outline of this work

II.11.1 Distribution of genetic variation and host specificity among nitrogen fixing symbionts from closely related *Lotus* species

In this study, nodule-associated bacteria of natural populations of two European *Lotus* species that represent two different rhizobia compatibility groups, *L. corniculatus* and *L. pedunculatus,* were identified and characterized. The host species were sampled at 21 locations in a north-south transect through Europe, resulting in a collection that comprises 71 individuals of *Lotus* and 41 bacterial strains. The early and late symbiotic phenotypes induced by the newly isolated bacterial strains were compared with those described for genetically modified rhizobial strains. The bacterial strains were tested on six *Lotus* species: *L. japonicus*, *L. corniculatus*, *L. filicaulis*, *L. burttii*, *L. glaber*, *L. pedunculatus* and three *L. japonicus* ecotypes: Gifu, MG20, Nepal. Variation was discovered among both hosts and symbionts in their ability to form symbiotic associations. The crossable host species and the newly identified microsymbionts with polymorphic symbiotic compatibilities are a valuable resource for the future identification of the genes involved in symbiotic compatibilities at different stages of the RNS.

II.11.2 Molecular evolution of Nod Factor Receptor 5 correlates with contrasting rhizobium specificity of European populations of *L. pedunculatus* and *L. corniculatus*

The levels of nucleotide variation were characterized at the *NFR5* locus*,* as one of the two loci critical in the recognition of the bacterial symbiont in natural European populations of two legumes species with contrasting specificity, namely *L. corniculatus* and *L. pedunculatus*. The *NFR5* locus was evaluated for evidence of natural selection. Regions likely to be involved in ligand recognition, i.e. the LysM containing ectodomains, were postulated to show signatures of positive selection. Intriguingly, sites that showed evidence of positive selection were found in the kinase domain and within the putative NF binding sites. We generated homology models of the LysM2 and the kinase domains of *NFR5* and superimposed the identified sites under positive selection onto these models. Furthermore, we explored the compatibility barriers in *L. japonicus, L. pedunculatus* and a F1 generation hybrid hereof by performing cross-inoculations with wild-type *M. loti* and *Bradyrhizobium*

sp. strains and a *M. loti* T3SS mutant.

II.11.3 <u>A quick T90 reporter line based assay to test the symbiotic activity of *L. japonicus* Gifu nodulation and AM mutants</u>

We tested T90 expression in the background of symbiotic mutant lines. The symbiotic GUS activity in these mutants was tested as a reporter for root symbioses in a transient root epidermis transformation assay. Double homozygous lines of T90 in the background of symbiotic mutant lines were generated and tested for their GUS activity patterns. Two preconditions towards the usage of these newly generated symbiosis mutant lines were investigated: the rhizobium specificity of the symbiotic GUS induction and the activation via the common *SYM* pathway. More specifically, we tested if the lines are still unresponsive to symbiotic stimuli including NF or rhizobia matching their parental mutant line phenotypes, and if they respond to signals acting downstream of the introduced parental mutant allele.

III MATERIALS AND METHODS

III.1 Plant materials

III.1.1 Natural *Lotus* spp. populations

We sampled *Lotus* species and their nodulating bacteria at 21 sites in a North-South transect through Europe (Table 2). At each site one to four whole plant samples were taken across the transect – totaling 71 individuals of *Lotus*. If plants were producing seed at the time of collection, seeds were sampled for subsequent inoculation experiments.

Our sampling of *Lotus* species for the population genetics study of the *NFR5* locus comprises 18 individuals from 11 natural European populations, and five diploid autogamous standard lines and agronomic cultivars (III.1.2, Table 9). Plant and sampling sites for all newly amplified *NFR5* alleles are listed in Table 9.

The diploid and self-compatible ecotypes *L. japonicus* Nepal line was collected in April 2001, near Kathmandu (Nepal) by Dr. B. N. Prasad, deposited with the NBRP Miyazaki germplasm unit and seeds were kindly provided by Makoto Hayashi. We generated the S4 generation by recurrent selfing.

Table 2. Sampling locations of European *Lotus* populations.

Site No.	Place	Country	GPS coordinates		collection date
1	Wendelstein	Germany			Aug 13, 2006
2	Wendelstein	Germany			Aug 13, 2006
3	Kreuzeck, Bergstation	Germany			Aug 13, 2006
4[a]	Brauneck, Lenggriess	Germany			Aug 20, 2006
5	service area Autohof Soltau, A7 exit 45	Germany	52°56´44,5"	09°52´50,0"	Aug 08, 2006
6[a]	lake Yngaren near Nyköping, south-west of Stockholm, E4, Sverige	Sweden	58°48´02,4"	16°39´46,1"	Aug 10, 2006
7[a]	Kårböle skans, service area ahead of Kårböle, Hälsingland, route 84 from Ljusdal to Sveg	Sweden	61°58´32,4"	15°20´01,8"	Aug 13, 2006
8	Vålådalen – Ottsjö, Jämtland, south of Åre, west of Östersund	Sweden	63°13´19,4"	13°09´51,8"	Aug 15, 2006
9[a]	Sjoa at E6 intersection 51 to Heidal, Gudbrandsdalen south of Otta	Norway	61°39´29,7"	09°34´52,3"	Aug 17, 2006
10[a]	Skammestein at route 51 northwest of Fagernes, intersection to E16	Norway	61°10´54,6"	08°57´54,5"	Aug 17, 2006
11	Maristuen at E16 between Fagernes and Lærdal, Fillefjell	Norway	61°06´31,8"	08°02´07,0"	Aug 18, 2006
12[a]	Noresund, Hallingdal, at route 7 from Hønefoss to Gol, northwest of Oslo	Norway	60°11´01,9"	09°37´31,1"	Aug 19, 2006
13[a]	E6 Oslo-Göteborg, exit 11 Fredrikstad, Norge	Norway	59°21´22,6"	10°50´36,7"	Aug 20, 2006
14	Tiers, 30-50 m west of ´Partissenhof´	Italy			Sept 08, 2006
15	Villnösser Tal	Italy			Sept 02, 2006
16	Lungau, Lessach, Purgger Hof	Austria			Aug 08, 2006
17	Lungau, Göriachtal, dam at east shore of the Göriach creek	Austria			Oct 08, 2006
18[a]	Oberstedten, forest Ockstadt, Obeliskenschneise	Germany	50°13'52"	08°33'44"	Nov 04, 2006
19[a]	Oberstedten, forest Ockstadt, König-Wilhelms-Weg	Germany	50°13'43"	08°33'10"	Nov 04, 2006
20[a]	Stierstadt, Steinbacher forest, Königsteiner Pfad	Germany	50°11'29"	08°33'57"	Nov 04, 2006
21[a]	Feldberg, Hegewiese	Germany	50°14'51"	08°28'35"	Nov 04, 2006
22	Bornholm, Raghammer	Denmark			June 08, 2007
23	Hagön Naturreservat, 2 km south of Halmstad	Sweden			June 06, 2007

[a] Population sites with successful rhizobia isolation and culture

III.1.2 Standard lines and agronomic cultivars

The following standard lines were used: *L. burttii* B-303, *L. filicaulis, L. japonicus* ecotypes 'Miyakojima' MG20, Gifu B-129, and 'Nepal'.

The following agronomic cultivars were used for inoculation experiments: two diploid, cross-pollinated, Uruguayan cultivars were used, *L. pedunculatus* LE306 and *L. glaber* cv. Herminia (PAS S.A. Montevideo Uruguay), kindly provided by Mónica Rebuffo of INIA "La Estanzuela", Uruguay and diploid cross-pollinated Uruguayan cultivars *L. pedunculatus* LE306 and *L. glaber* cv. Herminia (PAS S.A. Montevideo Uruguay), the tetraploid commercial cultivar *L. pedunculatus* cv. Maku, and tetraploid autogamous *L. subbiflorus* cv. El Rincón. Additionally in-vitro clones of F1 generation hybrid 6.2 from the original cross *L. pedunculatus* LE306 x *L. japonicus* Gifu were used for inoculation experiments.

For nodulation experiments with our new isolate of *R.* cf. *leguminosarum* 10.2.1 (hereafter *Rl*-10.2.1), used *Pisum sativum* cv. Sparkle was used.

III.1.3 <u>Generation of L. pedunculatus x L. japonicus hybrids</u>

L. pedunculatus LE306 was used as receiving plant and pollinated with pollen collected from *L. japonicus* Gifu. Due to an early abortion of the young embryo all developing seeds were subjected to an embryo rescuing procedure. The developing embryos were separated from the surrounding endosperm and transferred to plates. Our cooperation partner, Mónica I. Rebuffo (National Institute of Agricultural Research INIA "La Estanzuela", Ruta 50, km 11, C.P. 70000 Colonia, Uruguay) completed the embryo recuing and provided us with in-vitro clones of F1 generation hybrids. We tested 87 putative F1 hybrids with four co-dominant microsatellite markers distributed across three chromosomes; TM0304, TM0088, TM0793, and TM0637. For marker sequences and positions are listed at the Kazusa DNA Research Institute *Lotus*-online database (http://www.kazusa.or.jp/lotus/clonelist.html). We confirmed the heterozygous genotype of three F1 hybrids; 6.2, 6.99, and 6.100 and confirmed the equality of their nodulation phenotypes with the strains, *Bradyrhizobium* sp. NZP2309 and *Mesorhizobium* MAFF303099 (Fig. 14a).

III.1.4 <u>T90 reporter lines with symbiotic mutant line background</u>

III.1.4.1 Generation of double homozygous lines

We used the *L. japonicus* ecotype B-129 Gifu transgenic reporter line T90 (Webb *et al*., 2000) carrying the promoterless uidA (β-glucuronidase; GUS) reporter gene from *E. coli*. We performed reciprocal crosses (Gresshoff, 1997) with this T90 reporter line and the mutant lines *symrk-10* and *nfr1-1* (Table 3; Table 4). From the F2 population we identified double homozygous individuals via PCR, nod-minus phenotypes, and sequencing of the respective mutant allele. The PCR and sequencing markers are listed in Table 6. To test the presence of the T90-derived insertion of the GUS construct we used the primer pairs LjCBP1-fwd / LjCBP1-rev and LjCBP1-fwd / T90-rev to amplify the wild-type *Lj-Cbp1* gene (giving rise to a product of 650 bp, Fig. 24a) or the T90 *Lj-Cbp1* (giving rise to a 1.3 kbp product, Fig. 24a). Of the F2 individuals that were identified homozygous for the T90 construct, we selected all non-nodulating ones and confirmed their homozygous status at the respective mutant locus by amplifying and sequencing the gene region close to the mutation. Hetero-/and homozygosity were determined using the mutation scanner function of the Staden package release 1.6.0 beta 4 (Staden, 1996).

Table 3: Derived mutant lines of *L. japonicus* ecotype B-129 Gifu used in this study

Gene/ Allele	Progenitor/ Mutant line	Genomic mutation	Effect	Nod/ AM phenotype	Reference
nfr1-1 (*sym1-1*)[a]	282-118	C4973 to T	Q493 to stop	nod⁻/ AM⁺	(Schauser *et al.*, 1998; Radutoiu *et al.*, 2003; Sandal *et al.*, 2006),
symrk-10	SL1951-6	G4447 to A	D738 to N	nod⁻/ AM⁻	(Perry *et al.*, 2003; Markmann *et al.*, 2008)
castor-2 (*sym4-2*)[a]	EMS1749	G1057 to A	A264 to T	nod⁻/ AM⁻	(Imaizumi-Anraku *et al.*, 2005; Kistner *et al.*, 2005)
ccamk-2 (*sym15-2*)[a]	cac57.3	G5307 to T	E453 to stop	nod⁻/ n.d.	(Demchenko *et al.*, 2004; Kistner *et al.*, 2005; Tirichine *et al.*, 2006)

[a] previous name; nod⁻, no nodulation; AM⁺, AM symbiosis; AM⁻, no AM symbiosis; n.d. not determined.

III.1.4.2 Double homozygous lines used in this study

Table 4: Double homozygous lines from crosses between common *SYM* mutants and the reporter line T90

Cross (female x male parent)	Plant ID	Reference
castor-2 x T90	L2473	(Kistner *et al.*, 2005)
ccamk-2 x T90	L2724	(Kistner *et al.*, 2005)
nfr1-1 x T90	L3686, L3687, L3688	(this study)
T90 x *symrk-10*	L3675, L3676, L3678	(this study)

III.2 Generation of in-vitro clones

We generated in-vitro clones from the F1 generation hybrid 6. For this about 2 cm of the upper branch tips were clipped, surface sterilized, by incubation in 70% EtOH for 1 min, and subsequent rinsing once with dH₂O, followed by incubation in 3% NaOCl for 15 min, and subsequent rinsing three times with dH₂O. Then the branch tips were transferred to sterilized glass jars containing 300 ml sand-vermiculite mixture and 100 ml dH₂0. They were grown at the same conditions as inoculated seedlings until secondary roots developed.

III.3 Pollen grain germination

Pollen grains were collected from fully opened flowers and germinated as described by Saito *et al.* (2007). After incubation for 2 to 3 h at 28°C pollen growth was immediately inspected under inverted light microscope.

III.4 Hairy root transformation

Transgenic hairy roots were induced on *L. japonicus* ecotype B-129 Gifu derived transgenic reporter line T90, derived mutant lines (Table 3), and F3 generation

individuals of T90 lines with symbiotic mutant background (Table 4) using *Agrobacterium rhizogenes* strain AR1193 (Stougaard *et al.*, 1987) as described by Díaz *et al.* (2005). The *A. rhizogenes* strain AR1193 was carrying either the T265D construct of CCAMK (Yano *et al.*, 2008) inserted in the over expression vector pUB-GW-GFP (Maekawa *et al.*, 2008, hereafter pUB-GW-T265D-GFP) or an empty vector control (pUB-GW-EV-GFP), were the gateway cassette was excised with *Pvu*I.

III.5 Transient root transformation – vacuum infiltration

The *A. rhizogenes* strain AR1193 carrying the T265D construct of CCAMK (pUB-GW-T265D-GFP) or an empty vector control (pUB-GW-EV-GFP) were grown at 28°C on LB plates with the appropriate antibiotics (50µg/ml carbenicillin, 50µg/ml kanamycin, 50µg/ml rifampicin). Two day-old colonies were used to inoculate 15 ml LB liquid medium with antibiotics. The cultures were grown at 28°C and 200rpm. Bacteria were harvested by centrifugation for 5 min at 3,000 x g in mid- to late-exponential phase (OD_{600} between 10.0 to 20.0) (Campanoni *et al.*, 2007). Then they were washed once in vacuum-infiltration medium (VIM; Table 5) and resuspended in VIM to a final density of OD_{600} 0.5. Each well of a 96-wells DNA extraction plate was filled with 1 ml of the VIM-bacterial suspension or VIM without bacteria as negative control (mock infiltration). Four three-day old seedlings per well were submerged into the *Agrobacterium* suspension. The whole plate was transferred to a desiccator where vacuum was applied and held for 1 min after bubbles were rising to the surface. Then the vacuum was rapidly released and the vacuum-infiltration procedure was repeated once.

The whole plate containing the vacuum-infiltrated seedlings was incubated at 18°C in the dark overnight. Then the seedlings were transferred to co-cultivation medium plates (CM) and inoculated with a DsRed expressing derivative of strain *M. loti* MAFF303099 suspended in FP medium. To each root 20µl of the rhizobium suspension at a density of OD_{600} 0.05 was applied. The plates were incubated four days at 24°C in a light/dark regime while the bottom half of the plates was covered with black cardboard to prevent light exposure of the roots.

Table 5: Transient root transformation media

	Vacuum-infiltration medium (VIM)	Co-cultivation medium (CM)
100x PDM salts (Chabaud *et al.*, 1996)	10.0 ml	10.0 ml
PDM iron + vitamins (Chabaud *et al.*, 1996)	10.0 ml	10.0 ml
$CaCl_2$ x H_2O	0.2 g	0.2 g
Sucrose	10.0 g	10.0 g
Agar-agar (Sigma-Aldrich)	–	7.5 g
	pH 5.8 (adjusted with KOH)	
Acetosyringone[1]	0.1 ml	0.1 ml
10mM BAP[1] (Sigma-Aldrich)	1.5 ml	–
10mM NAA[1] (Sigma-Aldrich)	0.05 ml	–

[1] added to medium after autoclaving and cooling down to ca. 50°C

III.6 Bacterial materials

III.6.1 Isolation and cultivation of bacteria

For bacterial isolations from roots of natural populations of *L. corniculatus* and *L. pedunculatus*, one pink nodule per plant was carefully detached from the root and surface sterilized, by first vortexing in dH_2O, then incubating for 5 min in 0.1% SDS + 10% bleach, followed by incubation in 70% EtOH for 1 min, and subsequent rinsing three times with dH_2O. Then nodules were smashed with a pipette tip and suspended in 200 µl 20Q liquid medium (Röhm & Werner, 1991). 10 µl of dilutions (1:10, 1:100, 1:1000) and the final rinse as negative control were plated on 20Q solid medium and incubated at 28°C for 3-14 days. Single colonies were subjected to three rounds of single colony isolation on TY (Beringer, 1974) and on yeast extract mannitol (YEM) (Vincent, 1970). Prior to use in inoculation experiments, we isolated spontaneous antibiotic resistant mutant strains by plating the bacteria on agar plates with antibiotic gradients of 0-50 µg/ml (rifampicin, streptomycin, kanamycin, tetracycline) and 0-600 µg/ml (spectinomycin). Single colonies were subcultured on solid and liquid medium (120 rpm, 28°C) supplemented with the respective antibiotic concentration calculated from the position on the gradient plates. This resulted in 41 bacterial strains from these natural populations (Table 7). New rhizobium isolates of natural populations have been submitted to DSMZ and are being processed for deposit. We expect the DSM numbers shortly.

William J. Broughton (LBMPS, Université de Genève, Switzerland) kindly provided the GUS expressing NGR234 derivative. Additional reference rhizobium strains used for inoculations are all listed in Table 8.

III.6.2 Biparental mating

The constitutively GFP-expressing IncP plasmid pHC60 Tc^r (Cheng & Walker, 1998) was introduced via biparental mating from *E. coli* helper strain S17-1 Sm^r, Sp^r, Km^r (Simon *et al*., 1983) into a spontaneous antibiotic resistant (Sm^r, Rif^r) mutant of *Rl*-10.2.1 (Table 7). Donor strain S17-1.pHC60 was grown at 37°C in LB and recipient *Rl*-10.2.1 Sm^r, Rif^r was grown at 28°C in TY (Beringer, 1974). Both strains were grown to an OD_{600}: 0.6-0.8, harvested in 1:4 ratio (1.25 ml: 5 ml) by centrifugation, rinsed three times in PBS solution (137 mM NaCl, 2.7 mM KCl, 4.9 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7.4), resuspended in 30 µl TY, mixed and incubated for 3 h at 28°C, then plated on TY selective medium and incubated at 28°C for 3-7 days. Single colonies were subcultured and tested for their GFP fluorescence and stability of symbiotic phenotypes.

III.7 Bacterial inoculation experiments and phenotyping assay

III.7.1 Seed sterilization and germination

Seeds were surface sterilized with in a 10% (w/v) bleach, 0.1% (w/v) SDS solution for 10 min, rinsed four times with dH_2O and incubated at 4°C overnight. The seeds then were transferred to plates containing Fahraeus Plant (FP) medium (Fåhraeus, 1957) and grown for two to three days in the dark. Six-day-old seedlings and hairy root transformed *L. japonicus* plants were inoculated with rhizobium strains as described in (Stracke *et al*., 2002). Growth conditions were 24°C/8°C at 16-h-light/8-h-dark cycles.

III.7.2 Weck jar inoculation

Bacteria were applied at a final optical density (OD_{600}) of 0.01 in 500 ml glass jars (Weck, Wehr, Germany) containing 300 ml Seramis (Mars) and 150 ml liquid FP medium. For the wide nodulation assay the nodulation phenotypes of the plants were assessed at five weeks post inoculation with rating categories listed in Table 8.

III.7.3 Plate inoculations

GUS activity assays with T90 reporter line were performed on FP agar plates with bottom part covered with black cardboard. Each root was inoculated with 20 µl of bacterial solution (OD_{600}: 0.01). NFs of *Mesorhizobium loti* strain R7A at a concentration of 10^{-8} M were applied to roots of six-day-old seedlings of the reporter line T90 and T90 lines with symbiotic mutant background. Plants were grown for different time periods to test early and late nodulation phenotypes and GUS induction.

III.7.4 Co-inoculations

For co-inoculation experiments, rhizobium strains were grown individually to the exponential phase and inoculated in a 50:50 mixture at a final OD_{600} of 0.01. The following rhizobial strains were applied; derivatives of *M. loti* MAFF303099 expressing DsRed or β-galactosidase (*lacZ*) gene, and *lacZ* labeled type III secretion system deletion mutant DT3S (Okazaki *et al.*, 2010), and *Bradyrhizobium* sp. NZP2309. Nodulation phenotypes were assayed 5 weeks post inoculation.

III.7.5 X-Gal staining for rhizobial expression of *lacZ*

For the staining of *lacZ* activity in rhizobia, harvested roots were fixed by three round of vacuum infiltration with a PBS solution containing 2% glutaraldehyde and subsequent incubation for 1.5 h at 37°C. After washing twice with PBS roots were vacuum infiltrated with the staining solution (final concentration 0.8 mg/ml X-gal, 2.5 mM potassium ferrocyanide, 2.5 mM potassium ferricyanide in 0.2x PBS) and incubated at 15-30 min. When vascular bundles started to stain blue staining solution was exchanged to 60% EtOH and transferred to 4°C until phenotyping.

III.8 Histochemical localization of GUS activity in plant tissue

T90 reporter line plants were tested for GUS activity by incubation in staining solution (H_2O, 0.1 M $NaPO_4$ pH 7.0, 10.0 mM EDTA, 0.1% Triton X-100, 1.0 mM $K_3Fe(CN)_6$, 1.0 mM $Fe(CN)_6$, 1 mM 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-Gluc, Carl Roth, Karlsruhe, Germany) dissolved in N,N-Dimethylformamide at 37°C for 24 h after vacuum infiltration. Samples were examined under stereo light microscopy for blue staining.

III.9 Histochemical Analysis and Microscopy

Nodules were fixed in FAA solution (5% (v/v) acetic acid, 50% (v/v) EtOH, 3.7% (v/v) formaldehyde) dehydrated and embedded in Technovit 7100 resin (Heraeus Kulzer, Wertheim, Germany), and semi-thin sections (2-3 µm) were cut with a rotary microtome (RM2125 RT, Leica). Freshly harvested root material was embedded in 6% agarose blocks and sections (30-60 µm) were cut with the vibratome VT1000S (Leica, Bensheim, Germany). All sections were stained with toluidine blue (0.05%). Mircotome sections were mounted in DePex (Merck, Darmstadt, Germany), vibratome sections in 10% glycerol and immediately inspected under inverted light microscope.

III.10 DNA preparation, PCR and sequencing

Plant DNA was extracted from leaves using the QIAGEN DNeasy Plant Mini Kit (Hilden, Germany) following the manufacturers protocol. Rhizobial chromosomal DNA was extracted with MACHEREY-NAGEL NucleoSpin Plasmid kit (Düren, Germany) with extensive vortexing to shear the genomic DNA. The PCR and sequencing primers are listed in Table 6.

Table 6. Oligonucleotides used as PCR or sequencing primer

Primer name	Target gene	5'–3' nucleotide sequence
8f (Kunishima *et al.*, 2001)	16S rRNA	AGAGTTTGATCMTGGCTCAG
1492r (Kunishima *et al.*, 2001)	16S rRNA	GGYTACCTTGTTACGACTT
nodCF (Laguerre *et al.*, 2001)	nodC	AYGTHGTYGAYGACGGTTC
nodCI (Laguerre *et al.*, 2001)	*nodC*	CGYGACAGCCANTCKCTATTG
nodCfor540 (Han *et al.*, 2008)	nodC	TGATYGAYATGGARTAYTGGCT
nodCrev1160 (Han *et al.*, 2008)	*nodC*	CGYGACARCCARTCGCTRTTG
Nfr5shF	*NFR5*	GGATATTTTATTGACAATGTGAATGTTCC
Nfr5shR	*NFR5*	CTAGTTAAAAATGTAATAGTAACCACGC
Nfr5F2	*NFR5*	CCAAATCCAGCTAGGTGATAGC
Nfr5R2	*NFR5*	GCAACTCTATCAGAAGCACCC
LjCBP1 fwd	Lj-*CBP1*	GAGTGAGCACGCGCAAATAGC
LjCBP1 rev	Lj-*CBP1*	CTGAAGGGGTCTACTCTTGTGG
T90 rev	*uidA* (β-glu-curonidase; GUS)	CGGGATAGTCTGCCAGTTCAGTTC
Symrk 3-1F	*Lj-SYMRK*	CTACTTTCAGCAATACAGCATGAG
Symrk 3-2 rev	*Lj-SYMRK*	CTTTATGTTGAGAGGTTCCCGGCC
NFR1 3-3	*Lj-NFR1*	TGCTACAAGTGCTTAATTACGCTG
NFR1 3-4 rev	*Lj-NFR1*	CAAAGTACAAAACTTTTTGTCTGC

III.10.1 <u>Amplification and cloning of the 16S rRNA gene and *nodC*</u>

PCR fragments were cleaned with MACHEREY-NAGEL NucleoSpin Extract II kit (Düren, Germany) and directly sequenced with PCR primers. All sequences were deposited at GenBank (Table 7).

III.10.2 <u>Amplification and cloning of the NFR5 gene</u>

A fragment of 1,988 kbp of the *NFR5* gene was amplified using the primers Nfr5shF and Nfr5shR. PCR fragments were cleaned using MACHEREY-NAGEL NucleoSpin Extract II (Düren, Germany) and cloned into the pCR4Blunt-TOPO vector using Zero Blunt TOPO PCR cloning kit for sequencing (Invitrogen, Karlsruhe, Germany). For each cloned PCR fragment 55 clones were tested by colony PCR and from a minimum of 8 positively tested clones the plasmids were isolated using the QIAGEN QIAprep Spin Miniprep Kit (Hilden, Germany) and sequenced using the primers Nfr5F2 and Nfr5R2. All sequences were deposited at GenBank (Table 9).

III.11 DNA sequence analysis

III.11.1 Phylogenic analysis

Gene phylogenies were calculated on the basis of two bacterial genes, 16S rRNA and *nodC,* and one legume gene, *NFR5*. Nucleotide sequences were aligned with ClustalX v.2.0.3 (Thompson *et al*., 1997) and edited in MacClade v.4.08 (Maddison & Maddison, 2005). Phylogenetic analyses were completed using PAUP* v.4.0b10 (Swofford, 2002). The phylogenetic relationships between the sequences (*nodC* gene: 475 kb included, 16S rRNA gene: 1,306 kb included, *NFR5* gene: 1,806 kb) were determined using maximum parsimony (MP, with starting tree via stepwise addition, 500 random replicates, TBR branch swapping algorithm) and neighbor joining (NJ, HKY85, gamma = 0.5). These methods yielded similar topologies. Statistical reliability was tested using 1000 bootstrap replicates.

III.11.2 Plant *NFR5* gene phylogeny and population genetics analysis

Only the statistically weak or non-supported branches (bootstrap values below 50%) within the 'corniculatus' clade were not consistently recovered with both methods, MP and NJ. Unambiguous character changes (nucleotide substututions) were mapped onto one of 100 most parsimonious trees using McClade v.4.08 (D. R. Maddison & W. P. Maddison, 2005).

For 61 alleles the number of haplotypes and segregating sites, the nucleotide diversity, Pi (Nei, 1987) and the average nucleotide divergence using Jukes and Cantor correction (Jukes & Cantor, 1969), within the *NFR5* locus were calculated using DnaSP v.5 (Librado & Rozas, 2009). Additionally we performed a sliding window analysis and assessed departure from neutral evolution using Tajima's D statistic (Jukes & Cantor, 1969).

The tests for positive selection in *NFR5* and the reconstruction of ancestral nucleic acid sequences at nodes A-E in the phylogenetic tree (Fig. 19) were performed with PAML (Yang, 2007). The reconstructed sequences were aligned to *L. pedunculatus* and *Lj* Gifu (Fig. 17).

III.12 Homology modeling

The homology modeling was performed at the SWISS-MODEL workspace (Arnold *et al.*, 2006). Templates for the NFR5 kinase and LysM2 domains were identified separately using the gapped BLAST (Altschul *et al.*, 1997) query against the SWISS-MODEL Template Library. Congruent with earlier studies of LysM domains from NFR5 and the *Medicago* homolog NFP (Mulder *et al.*, 2006; Radutoiu *et al.*, 2007) we choose the nuclear magnetic resonance resolved structural information of *E. coli* membrane-bound lytic murein transglycosidase D (MltD, PDB ID: 1e0g; (Bateman & Bycroft, 2000) as template for homology model of LysM2 domain of *L. japonicus* Gifu.

Among the best BLAST results for the kinase domain was Interleukin-1 Receptor-Associated Kinase 4 (IRAK4, PDB ID: 2nryC) with 35% identities. Nevertheless, we choose the crystal structure of tomato Pto (PDB ID: pdb3hgk) (Dong *et al.*, 2009) with only 29% identities as a template, as it is a plant RLK involved in MAMP interaction.

The sequences were aligned in 'alignment mode'. The alignments were visually inspected and manually edited with the Swiss-PdbViewer, Deep-View (Guex & Peitsch, 1997). They were resubmitted in 'project mode' for model building. The NFR5 kinase is lacking the activation loop (p+1) causing a gap when aligned to Pto. Because of the large distance between the two amino acids adjacent to the alignment gap the connection between the two atoms could represent a false bond.

IV RESULTS

IV.1 Distribution of genetic variation and host specificity among nitrogen fixing symbionts from closely related *Lotus* species

IV.1.1 Natural European *Lotus*-rhizobium associations

We sampled *Lotus* plants and their nodulating bacteria in a North-South transect through Europe (Fig. 1, Table 2). The sampled individuals were named according to following scheme: 's.p.i', with 's' specifying the collection site, 'p' an individual plant from the population sampled at that site, and 'i' the respective bacterial isolate from a single nodule of that plant. Plants were identified to species based on morphology and ITS sequence analysis. Species determination using these two independent methods was congruent in all cases. 16S rRNA gene sequence analysis was used to identify the 41 bacterial isolates to the generic level. Our bacteria collection included the following genera: *Mesorhizobium, Rhizobium, Bradyrhizobium,* bradyrhizobial isolates clustering with *Rhodopseudomonas*, *Burkholderia* and *Curtobacterium* (Table 7, Fig. 2, Fig. 3). The plant collection comprises 71 *Lotus* individuals, of which 38 were identified as *L. corniculatus* and 33 as *L. pedunculatus* (Table 7). Phylogenetic analysis of the 16S sequences revealed that most of the bacteria isolated from these two European *Lotus* species belong to two major clades: the *Mesorhizobium* clade or the *Bradyrhizobium* clade (Fig. 2).

We typically found two types of naturally occurring associations: *Mesorhizobium* with *L. corniculatus* and *Bradyrhizobium* including additional bradyrhizobial isolates with *L. pedunculatus.* We never found *L. corniculatus* associated with any isolate from the *Bradyrhizobiaceae.* Neither we found *L. pedunculatus* associated with *Mesorhizobium.* The two species, *L. corniculatus* and *L. pedunculatus,* are adapted to different conditions and in natural populations they occur in different habitats (Nawrath, 2005) and thus they have access to different microfloras. *L. pedunculatus*, in contrast to *L. corniculatus*, is a wetland indicator and is highly tolerant of flooding (Justin & Amstrong, 1987). *L. corniculatus* is clearly stressed when grown in deoxygenated flooded medium and it is an indicator for poor soils (James & Crawford, 1998). Conformingly, we have never found them co-occurring at the same collection site. In agreement with their consistent host

preference, strains of *Mesorhizobium* show little variation in their 16S and *nodC* sequences. Three newly collected bacterial isolates (all from site 21) cluster together with *Rhodopseudomonas* sp. and formed a sister group to the *Bradyrhizobium* clade in the 16S rDNA tree. However, the 16S rRNA gene region is a unsuitable marker for the classification of rhizobia within the family Bradyrhizobiaceae and cannot discriminate among genera (Vinuesa *et al.*, 1998).

The *nodC* genes from these three isolates were sequence identical to those of some strains of *Bradyrhizobium* and therefore, these isolates cluster within the *Bradyrhizobium* clade in the *nodC* analysis (Fig. 3). Although the three isolates have the same *nodC* sequences, they differ in their nodulation capability on non-*L. pedunculatus* hosts. Of the *Bradyrhizobium* isolates, only the soybean symbiont *B. japonicum* USDA110 Δ132 was unable to induce nodules on *L. pedunculatus* and this isolate carried a distinctive *nodC* gene sequence (Fig. 3, Table 8). Likewise, although the two strains NGR234 and *S. meliloti* 1021 have nearly identical 16S sequences, their *nodC* sequences were distinct and these isolates differed greatly in their host specificity (Fig. 3, Table 8).

Under natural conditions, *L. pedunculatus* appeared to be more efficient in excluding ineffective rhizobia from nodulating their roots compared to *L. corniculatus*, because all rhizobia isolated in the natural setting from *L. pedunculatus* formed pink nodules on *L. pedunculatus* (Table 8). In contrast, some nodules from *L. corniculatus* were infected by bacterial strains that are close relatives of *R. etli, R. giardinii* and *R. leguminosarum* bv. *viciae.* These bacterial species are compatible symbionts of other legume species in the genera *Phaseolus* and *Vicia,* but not of *Lotus,* although some stains of *R. etli* can induce nodules on *L. japonicus*, but these senesce early (Banba *et al.*, 2001).
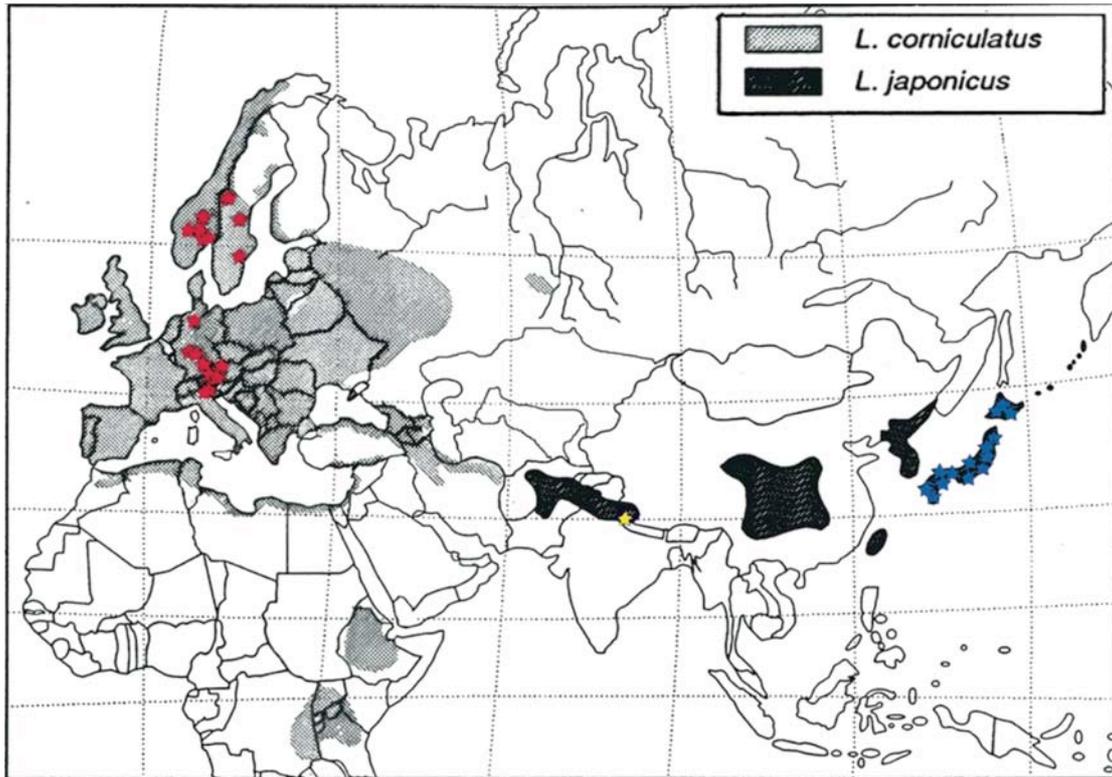
Figure 1. Sampling locations of *Lotus* species.
Sampling locations of *Lotus* species was superimposed onto distribution map of *L. corniculatus* and *L. japonicus* after Grant and Small (1996).Transect sampling locations of European *L. corniculatus* and *L. pedunculatus* (red stars); sampling site of *L. japonicus* ecotype Nepal (yellow star); original sampling locations of the Japanese ecotype collection of *L. japonicus* (blue stars).

Table 7. Bacterial isolates and Genbank accession numbers.

Strain / Collection ID[a]	Species	Host plant (origin)	16S rRNA accession no.	*nodC* accession no.
MAFF303099	*Mesorhizobium loti*	*Lotus corniculatus*	BA000012.4	BA000012.4
R7A	*Mesorhizobium loti*	*Lotus corniculatus*	HQ324666[b]	AL672113.1
ML105	*Mesorhizobium* sp.	*Lotus tenuis*	EU748910.1	EU748933.1
H152	*Rhizobium giardinii* bv. *giardinii*	*Phaseolus vulgaris*	NR_026059.1	AF217267.1
NGR234	*Rhizobium* sp.	*Lotus purpureus*	CP001389.1	U00090.2
1021	*Sinorhizobium meliloti*	*Medicago sativa*	AL591688.1	GQ507335.1
3841	*Rhizobium leguminosarum* bv. *viciae*	*Pisum, Vicia*	AM236080.1	AM236084.1
CIAT 652	*Rhizobium etli*	*Phaseolus vulgaris*	CP001074.1	CP001076.1
WM9	*Bradyrhizobium* sp.	*Lupinus* sp.	AF222751.1	AF222753.1
NZP2309	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324667[b]	HQ324654[b]
BC-C1	*Bradyrhizobium* sp.	*Chamaecytisus* sp.	AJ558030.1	AJ560654.1
USDA 110	*Bradyrhizobium japonicum*	*Glycine max*	BA000040.2	BA000040.2
ORS 1416ri	*Rhodopseudomonas* sp.	*Ononis natrix* subsp. *falcata*	AJ968691.1	
DUS751	*Burkholderia* sp.	*Mimosa pigra*	DQ316230.1	U386148.1
VKM Ac-2058	*Curtobacterium* sp.	perennial ice cover	AB042093.1	
wged11	*Leifsonia* sp.	ginseng root	DQ473536.1	
4.1.3[a]	*Rhizobium* sp.	*Lotus corniculatus* subsp. *alpinus*	HQ324673[b]	
6.1.1	*Rhizobium* cf. *leguminosarum*	*Lotus corniculatus*	HQ324671[b]	
6.1.2	*Rhizobium* cf. *leguminosarum*	*Lotus corniculatus*	HQ324672[b]	
7.1.1	*Burkholderia* sp.	*Lotus corniculatus*	HQ324703[b]	
7.1.2	*Burkholderia* sp.	*Lotus corniculatus*	HQ324702[b]	
7.1.4	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324679[b]	HQ324640[b]
7.2.1	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324680[b]	HQ324653[b]
9.4.1	*Rhizobium* sp.	*Lotus corniculatus*	HQ324668[b]	
9.4.2	*Rhizobium* sp.	*Lotus corniculatus*	HQ324669[b]	
9.4.3	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324682[b]	HQ324641[b]
9.4.4	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324683[b]	HQ324642[b]
10.1.2	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324676[b]	
10.2.1	*Rhizobium* cf. *leguminosarum*	*Lotus corniculatus*	HQ324670[b]	HQ324665[b]
10.2.2	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324677[b]	HQ324646[b]
10.2.3	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324678[b]	HQ324647[b]
12.4.1	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324685[b]	HQ324643[b]
12.4.2	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324686[b]	HQ324644[b]
13.1.1	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324681[b]	
13.1.2	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324674[b]	HQ324648[b]
13.1.3	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324675[b]	HQ324649[b]
13.2.1	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324704[b]	HQ324650[b]
13.2.2	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324705[b]	HQ324651[b]
13.2.3	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324706[b]	HQ324652[b]
13.3.1	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324684[b]	HQ324645[b]
13.3.2	*Mesorhizobium* sp.	*Lotus corniculatus*	HQ324687[b]	
18.1.1	*Curtobacterium* sp.	*Lotus pedunculatus*	HQ324707[b]	
18.2.1	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324691[b]	
19.1.1	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324693[b]	HQ324658[b]
19.3.1	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324690[b]	HQ324662[b]
19.4B.1	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324692[b]	HQ324659[b]
20.1A.2	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324698[b]	HQ324655[b]
20.2.1	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324688[b]	HQ324663[b]
20.5.1	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324695[b]	HQ324656[b]
21.1B.2	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324694[b]	HQ324660[b]
21.2A.1	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324696[b]	
21.2A.2	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324697[b]	HQ324664[b]
21.3B.1	bradyrhizobial isolate	*Lotus pedunculatus*	HQ324700[b]	
21.3B.2	bradyrhizobial isolate	*Lotus pedunculatus*	HQ324701[b]	HQ324657[b]
21.4C.1	*Bradyrhizobium* sp.	*Lotus pedunculatus*	HQ324699[b]	HQ324661[b]
21.4C.2	bradyrhizobial isolate	*Lotus pedunculatus*	HQ324689[b]	

[a] Collection nomenclature: site.plant.isolate; [b] New GenBank records

Table 8. Summary of late (5 wpi) nodulation phenotypes of *Lotus* species inoculated with rhizobia strains.

Bacterium species	Plant host (origin)	Strain / Collection ID (site.plant.isolate)	*Lotus japonicus* Gifu	MG20	Nepal	*Lotus filicaulis*	*Lotus burttii*	*Lotus pedunculatus*	*Lotus glaber*
Bradyrhizobium japonicum	*Glycine max*	USDA 110 Δ132 (Krause *et al.*, 2002)	Nod$^-$	Nod$^-$	Nod$^-$			bumps$^+$	bumps$^+$
Bradyrhizobium sp.	*Lotus pedunculatus*	NZP2309 (Bek *et al.*, 2010)	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^+$	
bradyrhizobial isolates		21.3B.2	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^+$Fix$^+$	Nod$^-$
		21.4C.2	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^+$	Nod$^+$Fix$^-$
Mesorhizobium loti	*Lotus corniculatus*	MAFF 303099 — DsRed[c]	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	bumps$^+$ Nod$^{(+)}$	Nod$^+$fix$^+$
		MAFF 303099 — Δ*nifH* (Ooki *et al.*, 2005)	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$		
		MAFF 303099 — *DT3S* (Okazaki *et al.*, 2010)	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$			bumps$^+$ / Nod$^+$Fix$^-$	Nod$^+$fix$^+$
Mesorhizobium sp.		10.2.2	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	bumps$^+$	Nod$^+$Fix$^+$
		13.1.2	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	Nod$^+$Fix$^+$	bumps$^+$	Nod$^+$Fix$^+$
Rhizobium cf. *giardinii*		4.1.3[b]	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$
R. cf. *etli*		9.4.1[a, b]	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$
R. cf. *leguminosarum*		6.1.1[b]	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	bumps$^+$	Nod$^-$
		10.2.1[a, b] GFP	Nod$^-$	bumps$^+$ / Nod$^{(+)}$Fix$^-$	bumps$^+$	Nod$^-$	bumps$^+$ Nod$^+$Fix$^-$	bumps$^+$	tumors$^+$
R. leguminosarum bv. *trifolii* cured of its own pSYM, provided with *R.l.* bv. *viciae pSYM* (pJB5JI)	*Vicia hirsuta, V. sativa, Trifolium subterraneum* (Spaink, HP *et al.*, 1987)	RBL5560 — *Rlv*-D (FITA *nodD*) (López-Lara *et al.*, 1996)	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$
		RBL5560 — *Rlv*-DZ (FITA *nodD*, *nodZ*) (Pacios Bras *et al.*, 2000)	Nod$^{(+)}$Fix$^-$	bumps$^+$	bumps$^+$ / Nod$^+$Fix$^-$	Nod$^-$	Nod$^+$Fix$^-$	bumps$^+$	Nod$^-$
		RBL5560 — *Rlv*-DZL (FITA *nodD*, *nodZ*, *nol*L) (Pacios Bras *et al.*, 2000)	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^+$Fix$^-$	Nod$^-$	Nod$^+$Fix$^-$	bumps$^+$ / Nod$^+$Fix$^-$	Nod$^+$fix$^-$
Sinorhizobium sp.	*Lablab purpureus*	NGR234 GFP (Schumpp *et al.*, 2009)	Nod$^+$Fix^0	Nod$^+$Fix^0	Nod$^+$Fix^0	Nod$^-$	Nod$^+$Fix^0	bumps$^+$	bumps$^+$
		NGRΩ*exoK* (Staehelin *et al.*, 2006)	Nod$^+$Fix^0	Nod$^+$Fix^0	Nod$^+$Fix$^+$	Nod$^-$	Nod$^+$Fix$^+$		
S. melitloti	*Medicago truncatula*	1021 RFP (Smit *et al.*, 2005)	Nod$^-$					Nod$^-$	Nod$^-$
Burkholderia sp.	soil / root associated	7.1.2[a]	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$
Curtobacterium sp.		18.1.1	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$	Nod$^-$		

[a] Co-isolated with *Mesorhizobium* from the same nodules; [b] Isolated from nodules of this host but non-effective symbiont; [c] kindly provided by M. Hayashi; nod$^-$: no organogenesis, bumps$^+$: small and flat swellings without cell differentiation, nod$^{(+)}$: small white uninfected nodules, nod$^+$: infected nodules with cell differentiation, tumors$^+$: extensively swollen epidermal and cortical cells; fix$^-$: plants with signs of nitrogen starving including chlorosis and growth similar to mock inoculated control, 'fix^0' plants without chlorosis but plant growth similar to mock inoculated control, fix$^+$: pink nodules and enhanced plant growth in comparison to mock inoculated control.
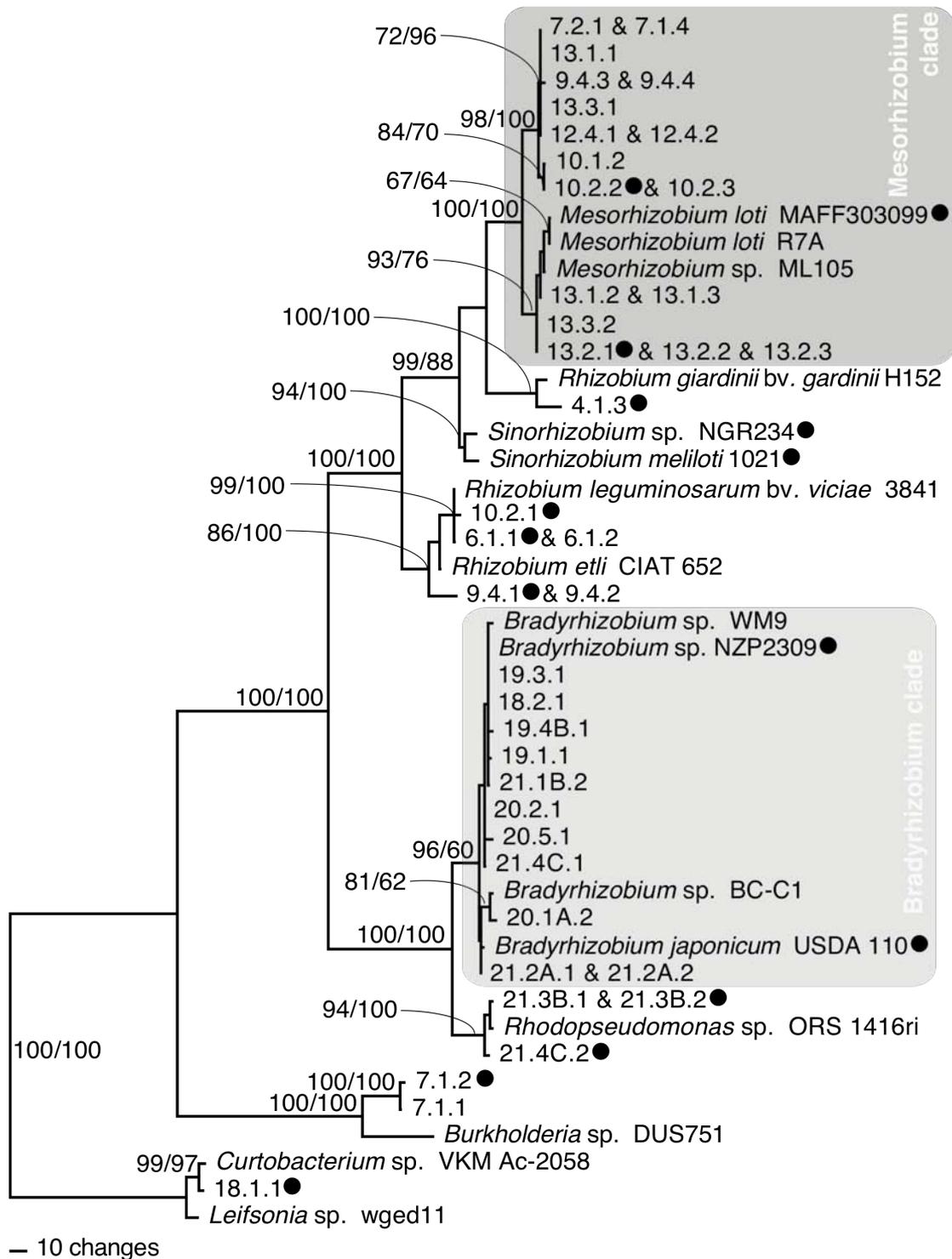
Figure 2. 16S rRNA gene phylogenetic tree.
Sequences of newly isolated bacterial strains from European *L. corniculatus* and *L. pedunculatus* populations and reference sequences were used to construct the tree under the maximum parsimony (MP) criterion. *Leifsonia* sp., *Curtobacterium* sp. and new isolate 18.1.1 were used as outgroup. MP and neighbor joining (NJ) bootstrap values higher than 60% are superimposed on the tree (MP/NJ). Strains tested for nodulation ability on three ecotypes of *L. japonicus* and three additional *Lotus* species (filled circles). Collection nomenclature of new isolates: site.plant.isolate.
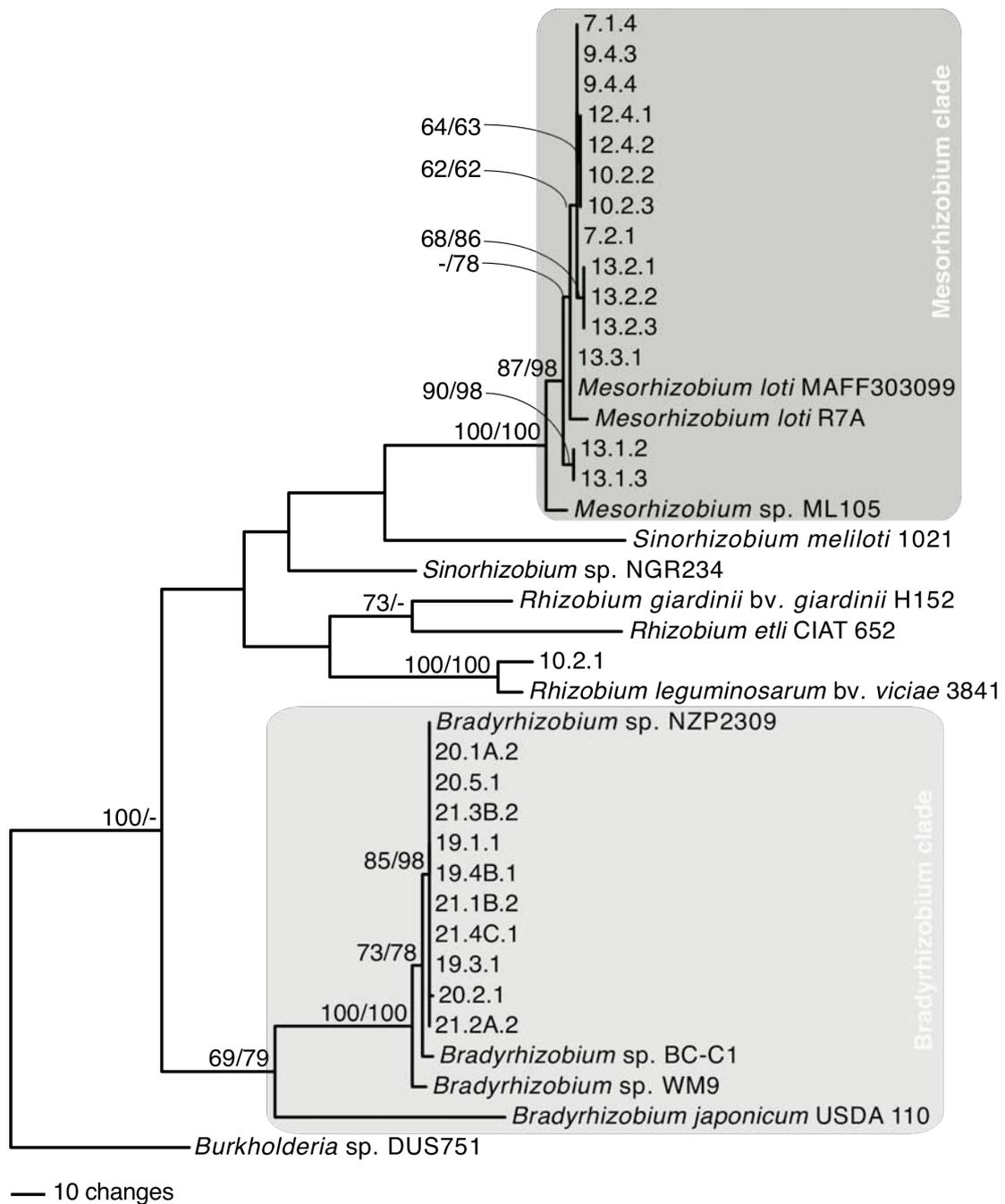
Figure 3. *nodC* gene phylogenetic tree.

Sequences of newly isolated bacterial strains from European *L. corniculatus* and *L. pedunculatus* populations and reference sequences were used to construct the phylogenetic tree under the maximum parsimony (MP) criterion. The tree was rooted with *Burkholderia* sp. sequence. One of the nine equally parsimonious trees shows a conflicting topology with *Bradyrhizobium* clade as sister clade to the *Mesorhizobium* clade. MP and neighbor joining (NJ) bootstrap values higher than 60% are superimposed on the tree (MP/NJ). Collection nomenclature of new isolates: site.plant.isolate.

46

In three cases, from a single nodule of *L. corniculatus* we isolated two different rhizobial strains. These double occupancies were from site 7 by isolates 7.1.4 (*Mesorhizobium* sp.) and 7.1.2 (*Burkholderia* sp.), from site 9 by 9.4.3/4 (*Mesorhizobium* sp.) and 9.4.1/2 (*Rhizobium* sp.) and from site 10 by 10.2.2/3 (*Mesorhizobium* sp.) and 10.2.1 (*Rhizobium* cf. *leguminosarum,* hereafter *Rl*-10.2.1) (Table 2; Table 7). Interestingly, all double occupancies occurred on *L. corniculatus* and included at least one *Mesorhizobium* sp. as rhizobial partner. In these associations, the *Mesorhizobium* strains were considered to be the compatible partners, substantiated by the ability of isolate 10.2.2 to form pink nodules on several *Lotus* species (Table 8). Only one of the non-*Mesorhizobium* co-isolated bacteria, strain *Rl*-10.2.1, was able to induce nodule formation on some *Lotus* species and on site 10 progeny. However these nodules were non-fixing. The other two strains from co-isolations, 9.4.1/2 (*Rhizobium* sp.) and 7.1.2 (*Burkholderia* sp.), failed to form nodules on any hosts tested and resembled mock-inoculations (Fig. 4).

Since *Burkholderia* includes species that are human pathogens, endosymbionts of AM fungi, soil bacteria, plant pathogens and nitrogen-fixing root nodule symbionts (Coenye & Vandamme, 2003). *Burkholderia* species were shown to be ancient symbionts of legumes, but there are no records for nodulation of *Lotus* species (Bontemps *et al.*, 2010). Our tested strain failed to induce any visible symbiotic and pathogenic symptoms on various species of *Lotus* (Fig. 4). We conclude that they probably represent saprophytic soil bacteria associated with plant roots. These naturally occurring double occupancies indicate that *Rhizobium* sp. 9.4.1/2, as well as *Burkholderia* sp. 7.1.2, are able to hitchhike along with effective symbionts.

Using standard primers and genus-specific degenerate primers, we failed to amplify *nodC* from isolates 7.1.1 and 7.1.2 (identified as *Burkholderia* sp.) and from 4.1.3, 6.1.1, 6.1.2, 9.4.1, 9.4.2 (identified as *Rhizobium* sp.). These isolates also failed to nodulate *Lotus* species. Only 6.1.1 was able to induce bumps on *L. pedunculatus*. Consistent with these phenotypic observations, failure to amplify the corresponding *nodC* sequences from these isolates could be due to sequence divergence at the *nodC* gene, due to the loss of symbiosis genes, or that these isolates never harbored the respective symbiosis genes. Sequences of 16S and *nodC* could be amplified from *Rl*-10.2.1, which induced nodule organogenesis on most tested *Lotus* species and this sequence is similar to that of *R. leguminosarum* bv. *viciae* 3841.
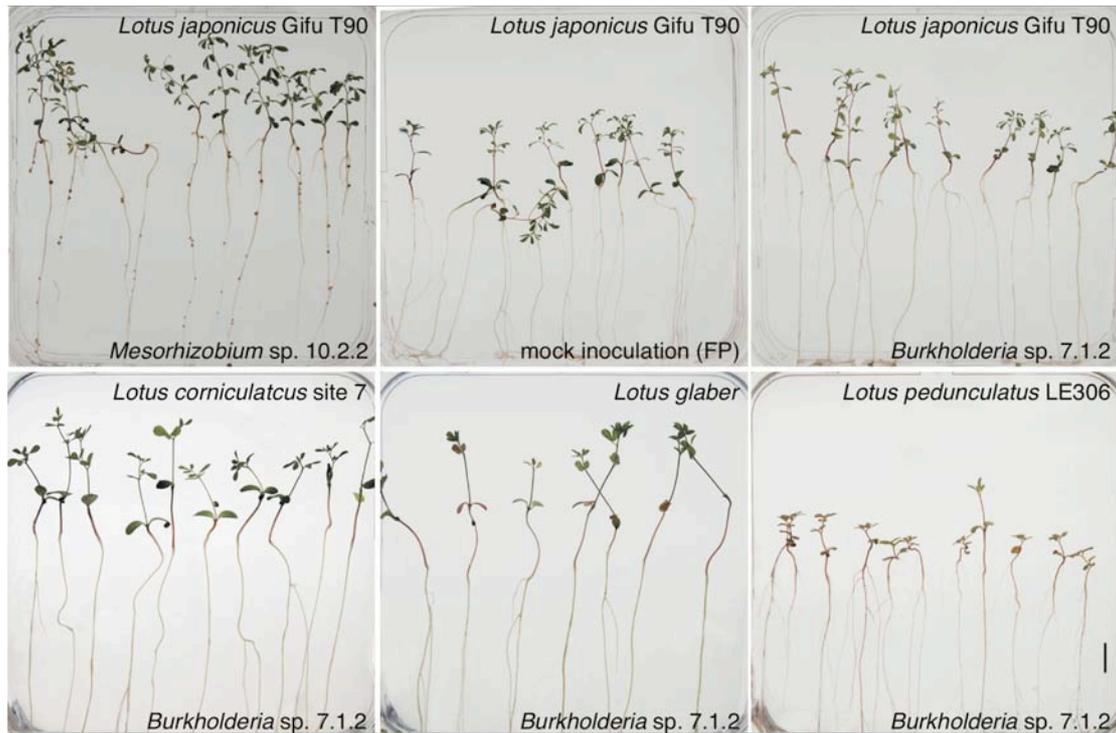
Figure 4. Plant phenotypes of various *Lotus* species 6 wpi.
Plant were inoculated on plates with new isolates *Burkholderia* sp. 7.1.2, *Mesorhizobium* sp. 10.2.2, and mock-inoculated control. Scale bar = 1 cm.

IV.1.2 Nodulation tests and RNS compatibility groups

We selected eight bacterial strains spanning the diversity across the 16S tree from our set of newly isolated bacteria and previously characterized strains and tested them for their nodulation ability on the progeny of host individuals from their isolation origin (Fig. 2). Additionally, these bacterial strains were inoculated onto three *L. japonicus* ecotypes and three other closely related *Lotus* species (Table 8). For the most part, our inoculation studies with the newly isolated strains of *Mesorhizobium* and *Bradyrhizobium* confirmed the two distinct compatibility groups reported earlier (Irisarri *et al.*, 1996; Saeki & Kouchi, 2000) (Table 8). Group 1 is the *Mesorhizobium* compatibility group, which can infect *L. japonicus, L. filicaulis, L. burttii, L. glaber,* and *L. corniculatus*. Group 2 is the *Bradyrhizobium* compatibility group, which can infect *L. pedunculatus.* For example, the *Mesorhizobium* sp. isolates 10.2.2 and 13.1.2 formed pink nodules when re-inoculated onto progeny from their original *L. corniculatus* host populations (sites 10 and 13, respectively). Likewise, the bradyrhizobial isolates 21.4C.2, 21.3B.2, and 19.4B.1 formed pink nodules on *L. pedunculatus* (Table 8). Specific exceptions to this general observation are described below.

Even though under natural conditions we failed to recover any bradyrhizobial strain from nodules of *L. corniculatus*, under experimental conditions, two bradyrhizobial strains (namely 21.4C.2 and NZP2309) could form non-fixing white nodules on *L. burttii, L. filicaulis, L. glaber* and *L. japonicus* (Table 8). The isolate 21.3B.2, as well as soybean symbiont USDA110 Δ132, failed to nodulate *L. japonicus* ecotypes or closely related species. Since the naturally occurring bradyrhizobial isolates have the same *nodC* gene and possibly the same nod genes sufficient for *L. pedunculatus* nodulation, an additional component is either absent or present leading to the Nod⁻ phenotype on *L. japonicus* and related species. *Mesorhizobium* sp. strains, although not isolated from natural populations of *L. pedunculatus*, could induce very small inefficient and non-infected nodules and bumps on *L. pedunculatus* LE306 (Table 8).

Cross-inoculations with the newly isolated *Rhizobium* strains led to a range of phenotypes on the different host species, ranging from Nod⁻ by 4.1.3, 9.4.1 on all hosts tested to the formation of bumps and nodules by 6.1.1 and 10.2.1 on some hosts. For example, although isolate 9.4.1 (*Rhizobium* sp.) failed to induce nodules or bumps on site 9 population progeny, isolate *Rl*-10.2.1, co-isolated with 10.2.2 (*Mesorhizobium* sp.), induced a small number of bumps on site 10 population progeny, indicating that both are not compatible for RNS with *L. corniculatus*.

IV.1.3 Early and late symbiotic capacity on *L. japonicus* Gifu T90 reporter line

To investigate early recognition of rhizobia prior to primordia and nodule formation, we inoculated the transgenic *L. japonicus* Gifu reporter line T90, which expresses β-glucuronidase under the control of a promoter approximately 1.3kb upstream of the *Cbp1* coding sequence (Webb *et al*., 2000). GUS activity in roots of T90 was previously found to be induced upon treatment with AM fungus *G. intraradices,* as well as with rhizobia *M. loti* NZP2037 and *M. loti* R7A NF, but not by a non-nodulating NodC⁻ mutant PN4047 (Kistner *et al*., 2005). This makes the T90 GUS activity a more sensitive parameter than e.g. nodulation to evidence macrosymbiont's sensing of a wide range of rhizobial signal molecules.

We found two GUS activity patterns of T90 upon inoculation with rhizobia. One pattern was characterized by a fast and strong GUS activation, which was typically observed with rhizobia capable of inducing complete or partial nodule organogenesis. The second pattern showed weak GUS activity that slowly increased

over time. This pattern was typically observed with rhizobia that did not induce organogenesis. All *Mesorhizobium* strains tested led to strong and rapid GUS induction on T90 with subsequent nodule formation and spatial contraction of GUS expression to the nodule zone (Fig. 5a,b). The bradyrhizobial strains NZP2309 and 21.4C.2 induced non-fixing nodules on T90 and showed the same rapid and strong GUS induction (Fig. 5a,b). *B. japonicum* USDA110 Δ132 induced strong GUS activity but failed to form nodules or bumps on T90 or on most *Lotus* species in the *Mesorhizobium* compatibility group (Fig. 5a,b; Table 8). Strong and rapid GUS activity was also induced upon *Rl*-10.2.1 inoculation, however nodules did not develop on T90 (Fig. 5a,b). Even the non-nodulating *Burkholderia* strain 7.1.2 led to GUS induction on T90 at a later time point, suggesting that bacterial signals other than compatible NFs have the ability to induce symbiotic GUS expression on T90 (Fig. 5a,b). We conclude that T90 GUS activity is a more sensitive parameter than nodulation for the detection of signaling molecules from bacteria.

To determine how variation in GUS activity and organogenesis depended upon NF structure, we used the genetically well-characterized strain of *R. leguminosarum* RBL5560. This is an isolate of *R. leguminosarum* bv. *trifolii* cured of its own Sym plasmid and provided with the Sym plasmid pJB5JI (= pRL1JI mep::Tn5) derived from *R. leguminosarum* bv. *viciae* (*Rlv*) (Beynon *et al.*, 1978; Wijffelman *et al.*, 1986; Spaink, HP *et al.*, 1987). In addition to this sym plasmid, this strain also carries the pMP1604 plasmid, which encodes a mutated *nodD604* gene (FITA *nodD*) and ensures the production of LCOs, independent of the presence of flavonoid inducers (López-Lara *et al.*, 1996). Three variants of this strain were used which differ in the NFs that they produce. RBL5560.pMP1604 (hereafter *Rlv*-D) produces LCOs, which are not modified at the reducing-terminal *N*-acetylglucosamine (GlcNAc). RBL5560.pMP1604.pMP2469 (hereafter *Rlv*-DZ) expresses FITA *nodD* and a fucosyltransferase (*nodZ*) from *B. japonicum* on pMP2469. This results in the production of LCOs with a fucosyl residue on C-6 of the reducing-terminal GlcNAc. RBL5560.pMP1604.pMP2469.pMP2470 (hereafter *Rlv*-DZL) carries pMP2470, leading to expression of an acetyltransferase (*nolL*) and resulting in the production of LCOs with 4-O-acetylation of the fucosyl residue (López-Lara *et al.*, 1996; Pacios Bras *et al.*, 2000).

The strain *Rlv*-D slowly induced weak GUS activity and no nodules on T90 (Fig. 5a,b). *Rlv*-DZ induced a fast and strong activity of GUS as well as

organogenesis (Fig. 5a,b). *Rlv*-DZL also induced fast and strong GUS activity as well as organogenesis on T90 (Fig. 5a,b). *Rlv*-DZ induced organogenesis to different degrees on various *L. japonicus* ecotypes while *Rlv*-DZL led to fully infected nodules (Fig. 6d; Table 8). These results are in agreement with earlier observations by Pacios-Bras *et al*., (2000) that the extend of compatibility between *L. japonicus* Gifu and the *Rlv*-D, -DZ and -DZL series is largely determined by the NF structure. In addition, our data demonstrate that the GUS induction pattern of the T90 line can be altered by structural modifications of the NF similar to the reported *NIN* induction (Rodpothong *et al.*, 2009).
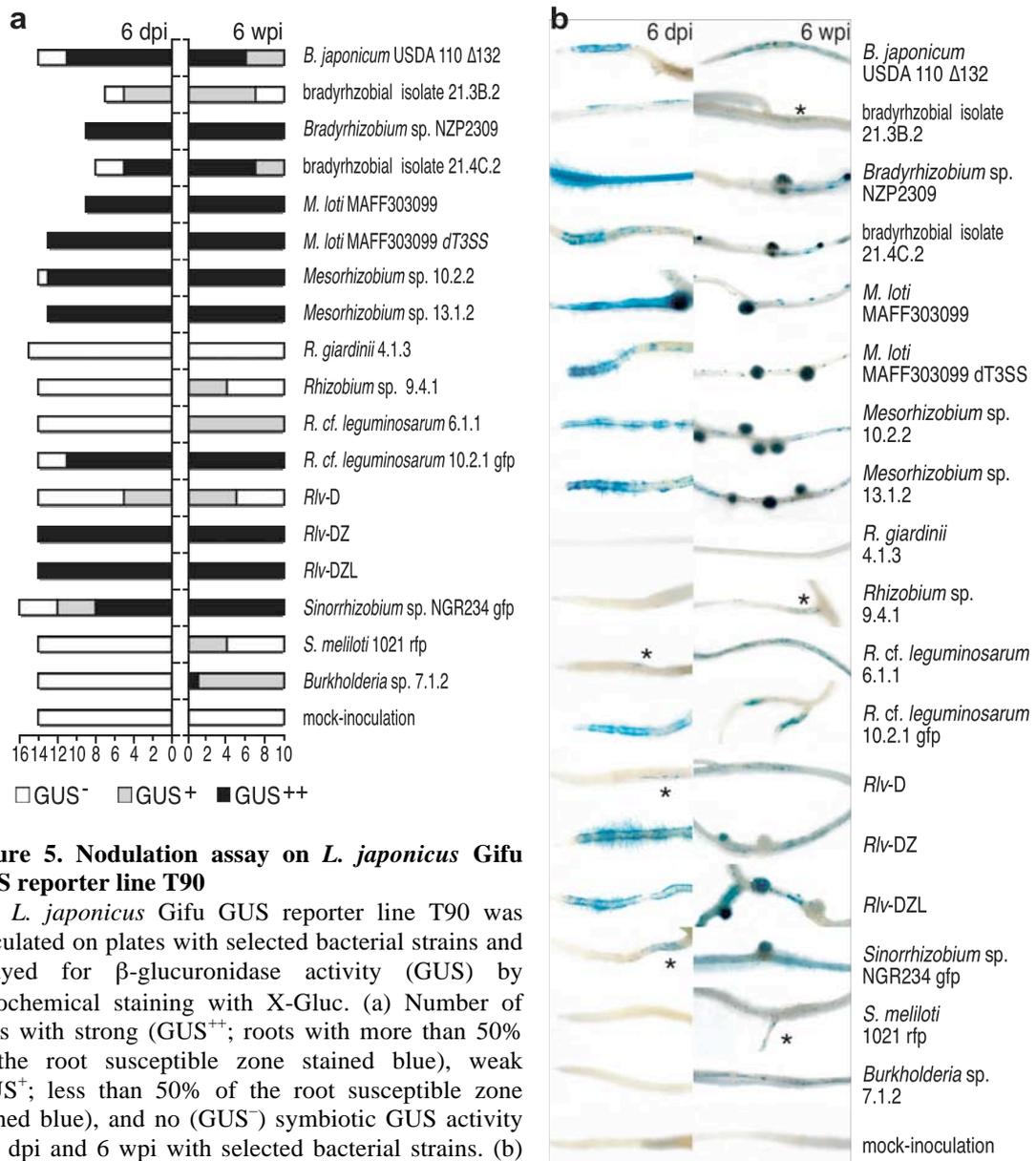
Figure 5. Nodulation assay on *L. japonicus* Gifu GUS reporter line T90

The *L. japonicus* Gifu GUS reporter line T90 was inoculated on plates with selected bacterial strains and assayed for β-glucuronidase activity (GUS) by histochemical staining with X-Gluc. (a) Number of roots with strong (GUS++; roots with more than 50% of the root susceptible zone stained blue), weak (GUS+; less than 50% of the root susceptible zone stained blue), and no (GUS−) symbiotic GUS activity at 6 dpi and 6 wpi with selected bacterial strains. (b) Early (6 dpi) and late (6 wpi) nodulation phenotypes and GUS activity induced by selected bacteria on reporter line T90. Blue staining corresponds to T90 GUS activity induced by the respective bacterial strain. Asterisk, GUS+-roots with a few blue stained cells only. Scale bars = 1 mm.

51

IV.1.4 <u>NF structure modulates organogenesis and infection phenotypes</u>

Among the newly isolated bacteria, strain *Rl*-10.2.1 was of particular interest because it provoked different meristem induction and different infection patterns upon inoculation of the *Lotus* species *L. burttii, L. filicaulis,* and *L. japonicus* including ecotypes MG20, Gifu, and Nepal (Table 8, Fig. 6a). To better characterize the infection process at the microscopic level, we generated a GFP-expressing derivative of this strain (Fig. 7). The infection process was arrested at different points depending on the host genotype: *Lj* Gifu and *L. filicaulis* were Nod⁻, *Lj* MG20 formed bumps and very small infected nodules, *Lj* Nepal showed zones with broadened root diameters and longitudinal stripes of GFP-expressing bacteria (hereafter 'elongated infection zones', EIZ) and *L. burttii* formed normal sized infected, but inefficient, nodules (Fig. 7a). Inefficiency of the symbiosis was evident from the nitrogen starving symptoms of the plant and high starch accumulation in non-infected cells (Fig. 8; Fig. 7e). A closer look at the morphology of an EIZ in *Lj* Nepal revealed longitudinal intercellular growth of *Rl*-10.2.1 rhizobia (Fig. 7c,d). Re-inoculation on progeny of *L. corniculatus* from the original collection site 10 led to bump formation only (Fig. 6a). On pea, *Rl*-10.2.1 formed wt ITs, but failed to form nodules even after 14 wpi (Fig. 7b). This reveals that *Rl*-10.2.1 can initiate root hair ITs, but that neither *L. japonicus* nor pea is a fully compatible host for this strain.
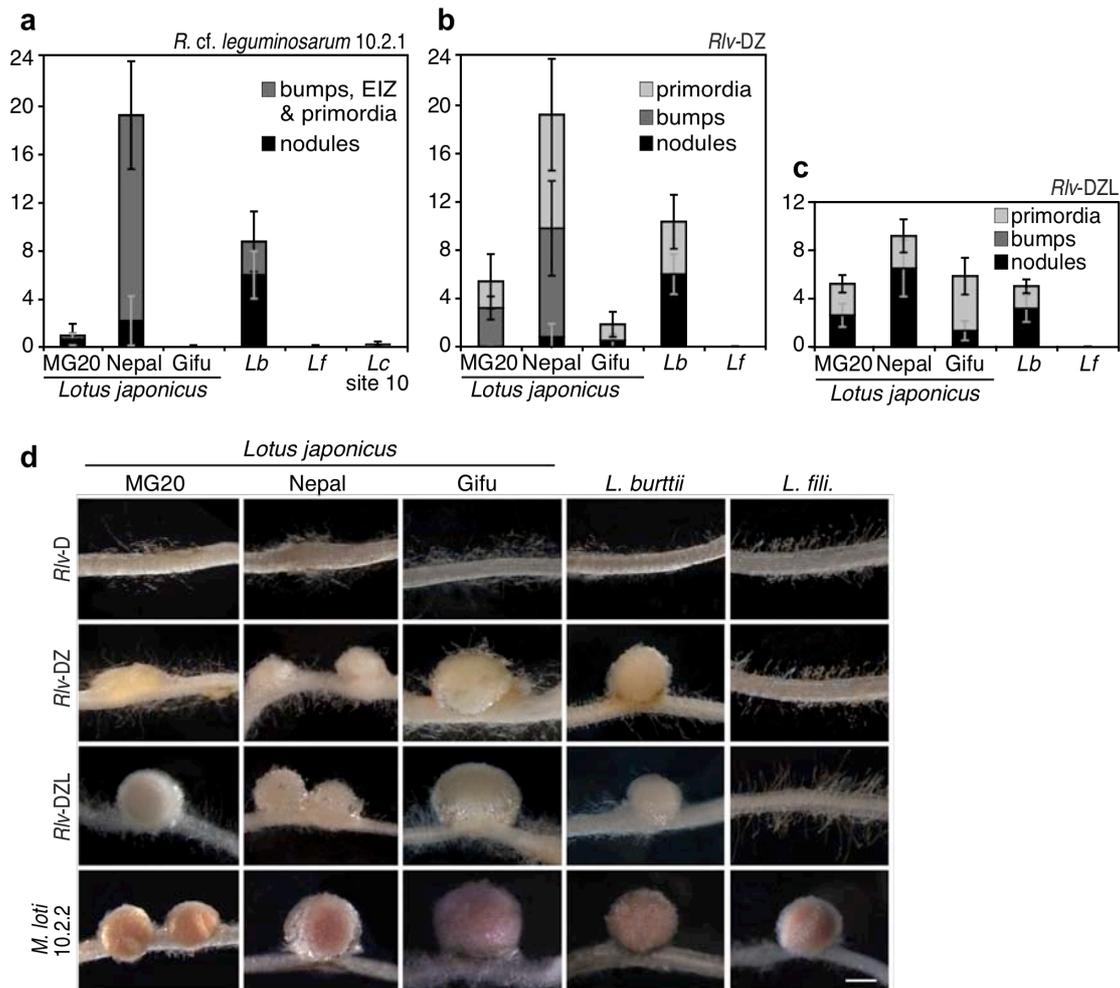
Figure 6. Variation of NF-dependent nodulation phenotypes on *Lotus* species.
(a to c), mean numbers of nodules, bumps, elongated infection zones (EIZ), and primordia on *Lotus* roots 6 wpi with (a), isolate *Rhizobium* cf. *leguminosarum* 10.2.1, and (b and c), derivatives of *R. leguminosarum* RBL5560 that produce structurally different NFs. *Lb*, *L. burttii*; *Lf, L. filicaulis; Lc, L. corniculatus* progeny of host population from site 10. Mean values are means of 10 plants. Error bars indicate standard deviations. (d), root nodulation phenotypes 6 wpi with *M. loti* 10.2.2 and three derivatives of *R. leguminosarum* RBL5560 that produce NF with different structures. Scale bar = 200 µm.
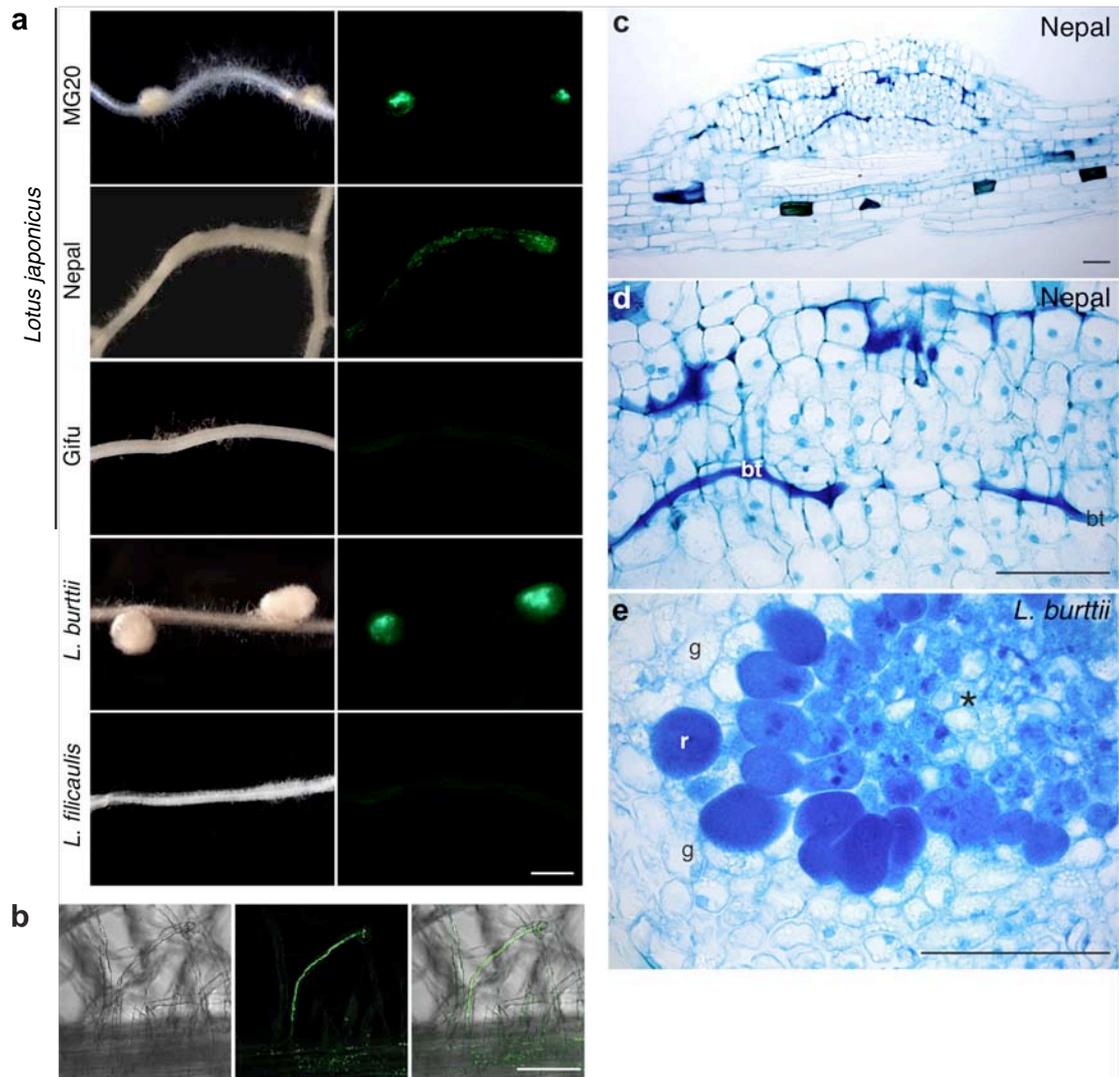
Figure 7. Nodulation phenotypes induces by strain *Rhizobium* cf. *leguminosarum* 10.2.1
(a), root nodulation phenotypes of various *Lotus* species 8 wpi with GFP-expressing derivative of *Rl*-10.2.1 under white light (left column) and infections with rhizobia expressing GFP (right column). (b), convocal micrographs of *Pisum sativum* SPARKLE infection threads induced by *Rl*-10.2.1 labeled with GFP. (c to e), light micrographs of microtome sections through infected *Lotus* roots at 6 wpi, stained with toluidine blue. (c and d), *L. japonicus* Nepal elongated infection zones; (c), overview, (d), close-up. (e), *L. burttii* nodule. bt, Intercellular bacterial threads; asterisk, central zone of the nodules with uninfected cells or early degrading bacteroids; g, cells filled with starch granules; r, and rhizobia. Scale bars = 1000 µm (a), 100 µm (b), 200 µm (c to e).
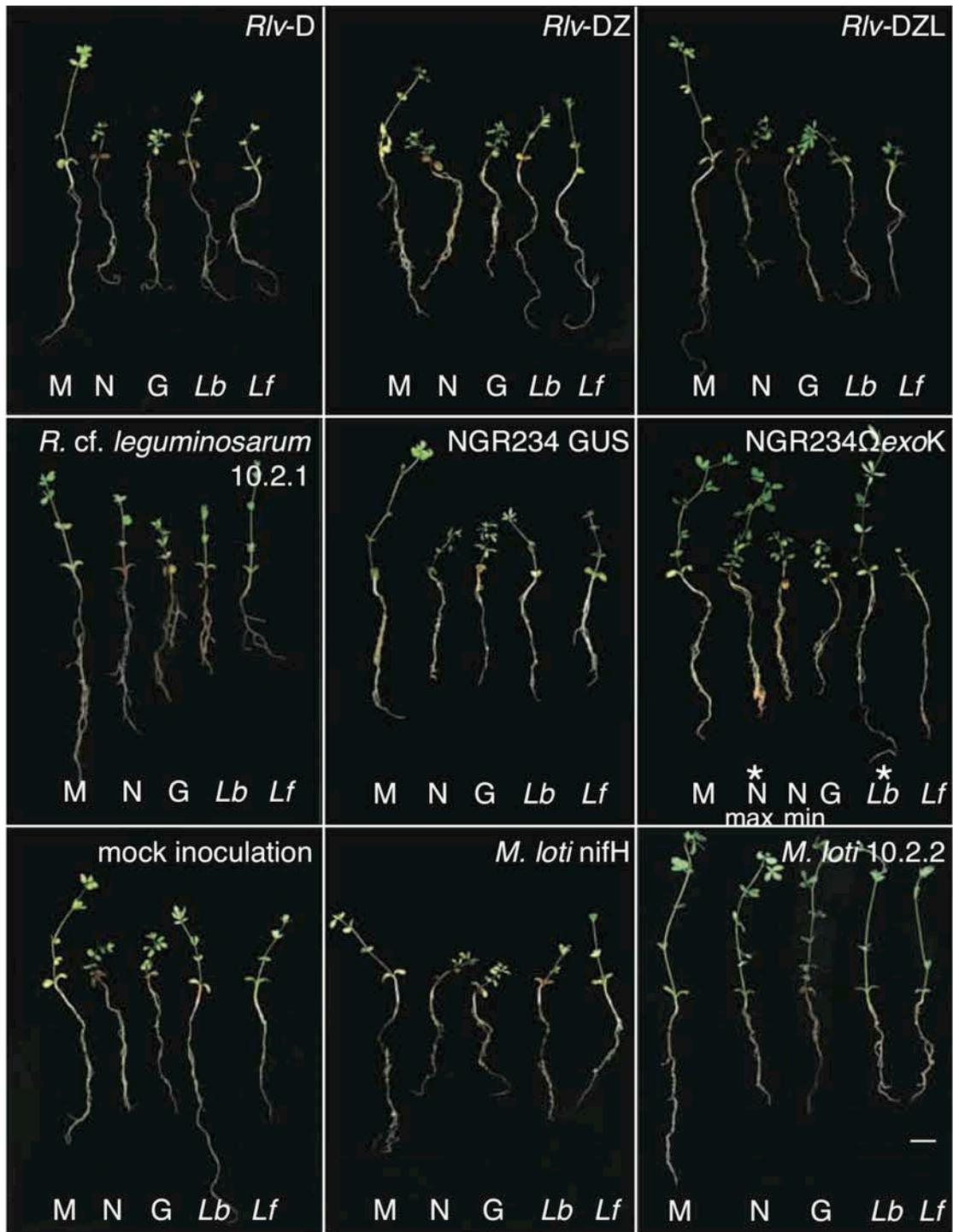
Figure 8. Growth differences of Lotus species 6 wpi with selected rhizobia strains and mock-inoculated control.
M *, L. japonicus* MG20*;* N *, L. japonicus* Nepal; G, *L. japonicus* Gifu; *Lb*, *L. burttii*; *Lf, L. filicaulis*; asterisk, plants with obvious growth enhancement. Scale bar = 1cm.
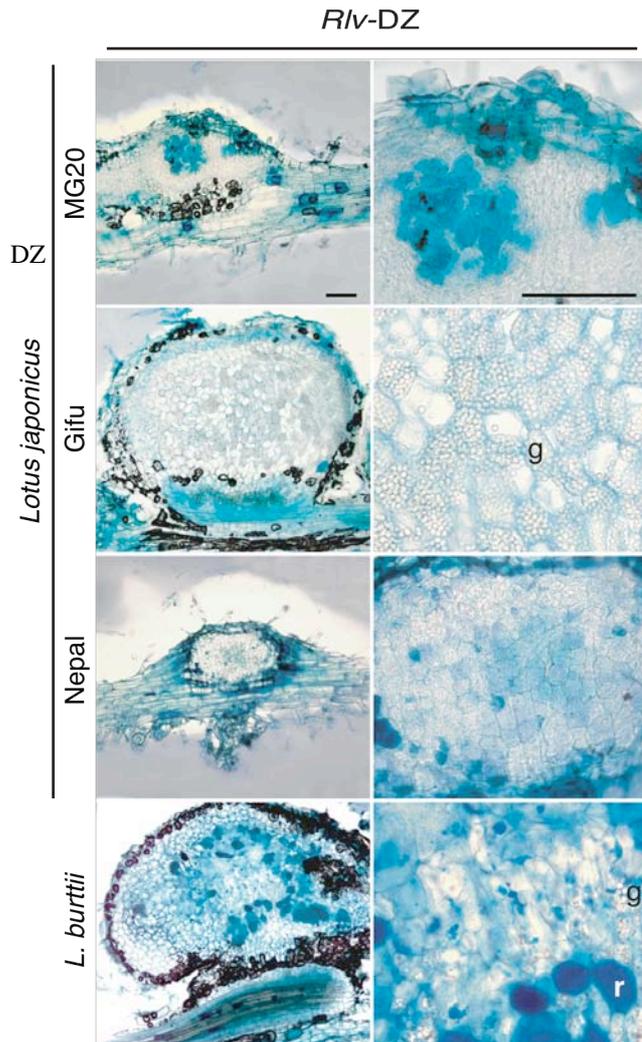
Figure 9. Light micrographs of nodules and bumps 6 wpi with Rlv-DZ.
Toluidine blue stained vibratome sections were taken through nodules or bumps from various *Lotus* species. *Rlv*-induced no organogenesis on *L. filicaulis*. Left column, overview; right column, close-up; g, cells filled with starch granules; r, and rhizobia. Scale bar = 100 μm.

To determine whether the differential infection phenotypes were specific to *Rl*-10.2.1, we compared the infection process across *Lotus* species by *Rl*-10.2.1 to that of derivatives of RBL5560. With *Rl*-10.2.1 and RBL5560 derivatives all hosts remained short and appeared to be nitrogen starved at 6 wpi indicating non-efficient symbiosis (Fig. 8). Nodules were not formed on any *Lotus* species upon infection by *Rlv*-D (Fig. 6d). In contrast, *Rlv*-DZ induced organogenesis on all *Lotus* species tested, excluding *L. filicaulis* (Fig. 6d). However, organogenesis was not complete at 6 wpi and led to a range of phenotypes: e.g. small white bumps on the three *L. japonicus* ecotypes and nodules on *L. burttii* (Fig. 6b). Cross sections through one of the few nodules formed on *Lj* Gifu revealed that the plant cells were not infected and were filled with many starch granules (Fig. 9). Infection by *Rlv*-DZL resulted in white nodules on all *Lotus* species tested, except for *L. filicaulis* (Fig. 6c,d). Quantitative analysis of these incomplete organogenesis phenotypes revealed similarities between *Rl*-10.2.1 and *Rlv*-DZ (Fig. 6a,b). For example, a similar proportion of complete (nodule) and

incomplete (primordia, bumps) organogenesis in response to inoculation by *Rl*-10.2.1 and *Rlv*-DZ was found on *Lj* Nepal and *L. burttii*. However, EIZ on *Lj* Nepal were only formed by *Rl*-10.2.1 (Fig. 7a,c,d). Interestingly, *Rlv*-DZL induced complete organogenesis on *Lj* ecotypes, but resulted in a reduced number of nodules on *L. burttii* (Fig. 6c).

The nodules formed by *Rlv*-DZL showed a central zone of enlarged plant cells filled with rhizobia (Fig. 10b). In contrast, *Rlv*-DZ showed infection incompatibilities with enlarged ITs and partial uptake of the rhizobia into plant cells, but without differentiation of the host cells (ecotypes MG20 and Nepal) (Fig. 10a). Only on *L. burttii* did nodules form with a central zone of differentiated cells filled with rhizobia and proper vascularization. These results suggest that early organogenesis (formation of primordia or bumps) requires a definite, but lower, stringency for NF structure (e.g. compare *Rlv*-D to *Rlv*-DZ), than later steps of organogenesis including differentiation of nodule cells coupled with infection of cortical cells, both of which require more stringent NF structures.
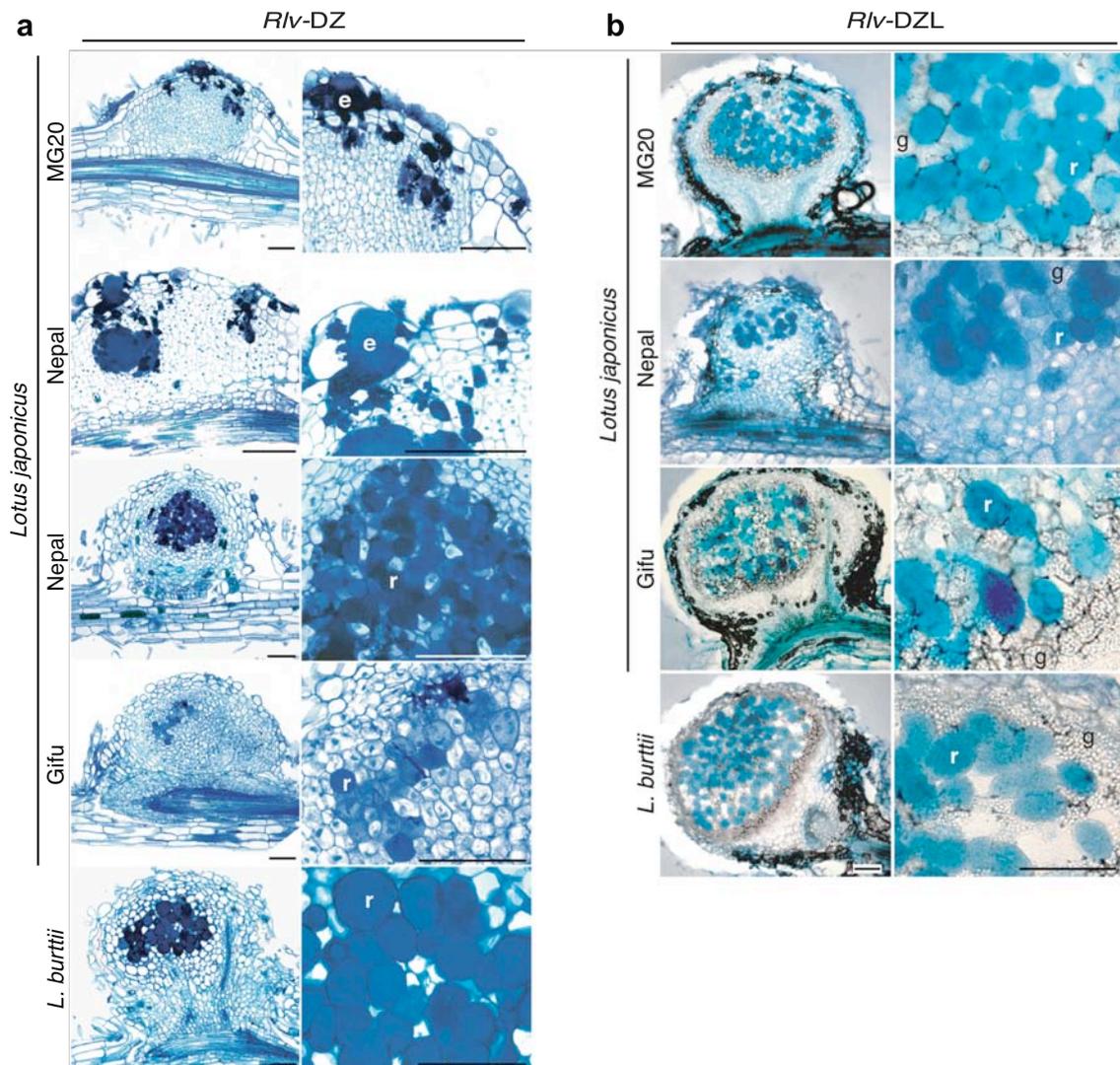
Figure 10. Variation of NF dependent infection phenotypes on *Lotus* species.
Light micrographs of toluidine blue stained sections through nodules and bumps induced 6 wpi by two derivatives of *R. leguminosarum* RBL5560 on various *Lotus* species. Both rhizobium derivatives fail to induce organogenesis on *L. filicaulis*. (a), microtome sectioned nodules and bumps induced by *Rlv*-DZ (b), vibratome sectioned nodules induced by *Rlv*-DZL. Left column, overview; right column, close-up; e, enlarged infection thread; g, cells filled with starch granules; r, cells filled with rhizobia. Scale bars = 100µm.

IV.1.5 Post-organogenesis incompatibility determinants

We identified variation among *Lotus* species in both early and late stages of the symbiosis with NGR234. A GUS expressing derivative of NGR234 induced the formation of ITs on all *Lotus* species tested (Fig. 11a). On *L. filicaulis* IT growth was arrested early in the root hair. Therefore, infected nodules formed on all *Lotus* species except on *L. filicaulis* (Fig. 11b). NGR234, even though not as efficient as *M. loti*, formed nodules on *L. japonicus*, and fixed sufficient nitrogen to avoid yellowing of the plant. *Lj* MG20 plants inoculated with NGR234 were reported to be clearly larger than mock-inoculated controls at eight wpi (Schumpp *et al*., 2009). Under our

58

conditions, we did not see a clear growth difference on plants six weeks following infection with NGR234 compared to mock-inoculations. NGRΩ*exoK*, which carries a mutation in *exoK* that encodes a putative glycanase, produces EPS devoid of low-molecular-weight forms (Staehelin *et al.*, 2006). Earlier studies showed that the absence of EOS caused incompatibility phenotypes on some hosts (e.g. *Albizia lebbeck* and *Leucaena leucocephala*), but on other hosts (e.g. *Vigna unguiculata*), nitrogen-fixing nodules were formed. Similar to the wild-type, NGRΩ*exoK* induced nodules on all tested *Lotus* species, except on *L. filicaulis* (Fig. 11b). On *L. burttii* and some *Lj* Nepal plants, NGRΩ*exoK* induced enhanced plant growth and increased fresh weight compared to the wild-type (Fig. 8; Fig. 12, asterisks). This suggests a negative role on the nodulation process by EOSs post organogenesis. The growth enhancement due to the lack of EOS production by NGRΩ*exoK* may be caused by accelerated infection or uptake of bacteria.
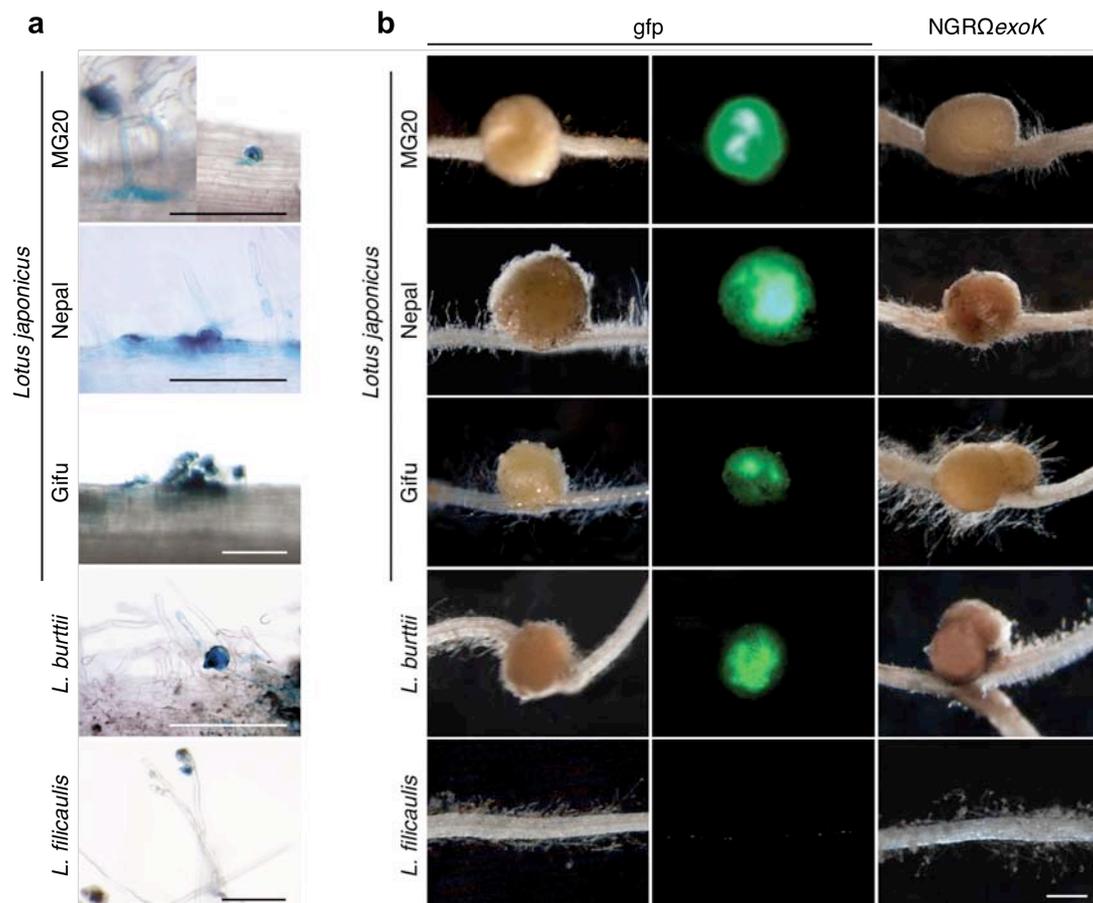
Figure 11. Early infection and late nodulation phenotypes induced by NGR234 and NGRΩ*exoK*
(a), root hair infection at 11 dpi with GUS expressing derivative of NGR234. (b), nodulation phenotypes at 6 wpi with GFP-expressing derivative of NGR234 under white light (left column) and infections with rhizobia expressing GFP (middle column) and nodulation phenotypes with exopolysaccharide mutant NGRΩ*exoK* under white light (right column). Scale bar = 200 µm (a), 500 µm (b).
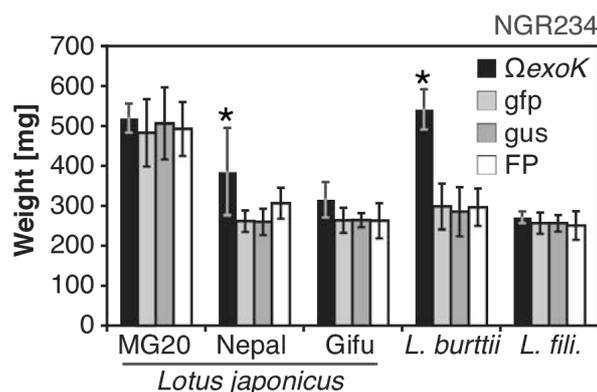
Figure 12. NGRΩexoK promotes plant growth on *L. burttii* and *L. japonicus* Nepal.
Various *Lotus* species were inoculated with derivatives of NGR234 expressing GFP or GUS, and with exopolysaccharides mutant NGRΩ*exoK*. Fresh weight was measured 6 wpi. Mean values are means of 10 plants. Error bars indicate standard deviations.

IV.2 Molecular evolution of Nod Factor Receptor 5 correlates with contrasting rhizobium specificity of European populations of *L. pedunculatus* and *L. corniculatus*

IV.2.1 Symbiotic compatibility is controlled at two levels

We tested the symbiotic compatibility of *Lj* Gifu and two lines of *L. pedunculatus* (LE306 and cultivar Maku) with *M. loti* MAFF303099 (hereafter *Meso*MAFF303099) and *Bradyrhizobium* sp. NZP2309 (hereafter *Brady*NZP2309) (Fig. 13; Fig. 14b). *Lj* Gifu formed efficient nodules with *Meso*MAFF303099 and *L. pedunculatus* formed efficient nodules with *Brady*NZP2309 (Fig. 13; Fig. 14b). *Meso*MAFF303099 was incompatible on *L. pedunculatus* and *Brady*NZP2309 was incompatible on *Lj* Gifu. However, incompatibility arose at different points depending on the cross-inoculation combination. Incompatibility between *Meso*MAFF303099 and *L. pedunculatus* occurred at the stage of infection. Here the rhizobia induced only a low number of ITs, which were aborted before reaching the base of the root hair cell (Fig. 13b). As a consequence, only non-infected bumps and small white nodules formed (Fig. 13a). In contrast, *Brady*NZP2309 was initially compatible with *Lj* Gifu and induced infected nodules, but these were white and inefficient (Fig. 14b). Inefficiency was deduced from the nitrogen starving phenotype of the plant.

To determine the dominance relationship of these incompatibility phenotypes, we inoculated in-vitro-clones of nine F1 generation hybrids from a cross between *L. pedunculatus* LE306 x *Lj* Gifu with *Meso*MAFF303099 and *Brady*NZP2309 (Fig. 14a). The hybrid lines formed fully efficient nodules with *Brady*NZP2309, but they formed almost exclusively white nodules and bumps when challenged with *Meso*MAFF303099. A more detailed phenotypic analysis was performed using the

60

hybrid 6.2 (hereafter *LpLj*6.2), for which we confirmed the heterozygous genotype. *Meso*MAFF303099 induced organogenesis and ITs on *LpLj*6.2 in-vitro-clones (Fig. 13a,b). The ITs progressed to the cortex of the nodule primordia, where they branched off without releasing the rhizobia into the host cells (Fig. 13b). Since IT progression in the hybrid *LpLj*6.2 was similar to the *L. japonicus* parent, initial compatibility appears to be a dominant trait encoded by *L. japonicus* alleles. However, later during the infection process, the rhizobia failed to be released into the host cells. This late infection blockage seems to be a dominant-negative trait inherited from the *L. pedunculatus.* When inoculated with *Brady*NZP2309, fully efficient pink nodules formed on the hybrid *LpLj*6.2 (Fig. 14a,b). Since this resembles the *L. pedunculatus* parent, the late efficiency incompatibility observed in *L. japonicus* to *Brady*NZP2309 appears to be recessive.

The segregation of these traits could not be assessed due to male sterility in the hybrids. All tested F1 hybrids produced degenerate pollen (Fig. 14c). Our attempts to obtain F2 hybrids or backcrossed progeny using Gifu as the male parent failed.

_____

Figure 13. Contrasting symbiotic phenotypes induced on *L. pedunculatus* LE306, *L. japonicus* Gifu and F1 generation hybrids upon inoculation with DsRed-labeled *M. loti* strain MAFF303099
(a), nodulation phenotypes at 5 wpi. (b), infection phenotypes at 1wpi.
*LpLj*6.2, F1 hybrid 6.2 from an original cross between *L. pedunculatus* LE306 x *L. japonicus*; bright field, white light micrographs; RFP, red fluorescence micrographs; overlay, overlay of both white light and fluorescence micrographs visualizing infections with DsRed-labeled rhizobia. (b) third column, hand section through one bump displayed above (a). Scale bars = 500 μm (a), 50 μm (b).

a

	L. japonicus Gifu	*L. pedunculatus*	F1 hybrid *LpLj*6.2	

b

	L. japonicus Gifu	*L. pedunculatus*	F1 hybrid *LpLj*6.2	
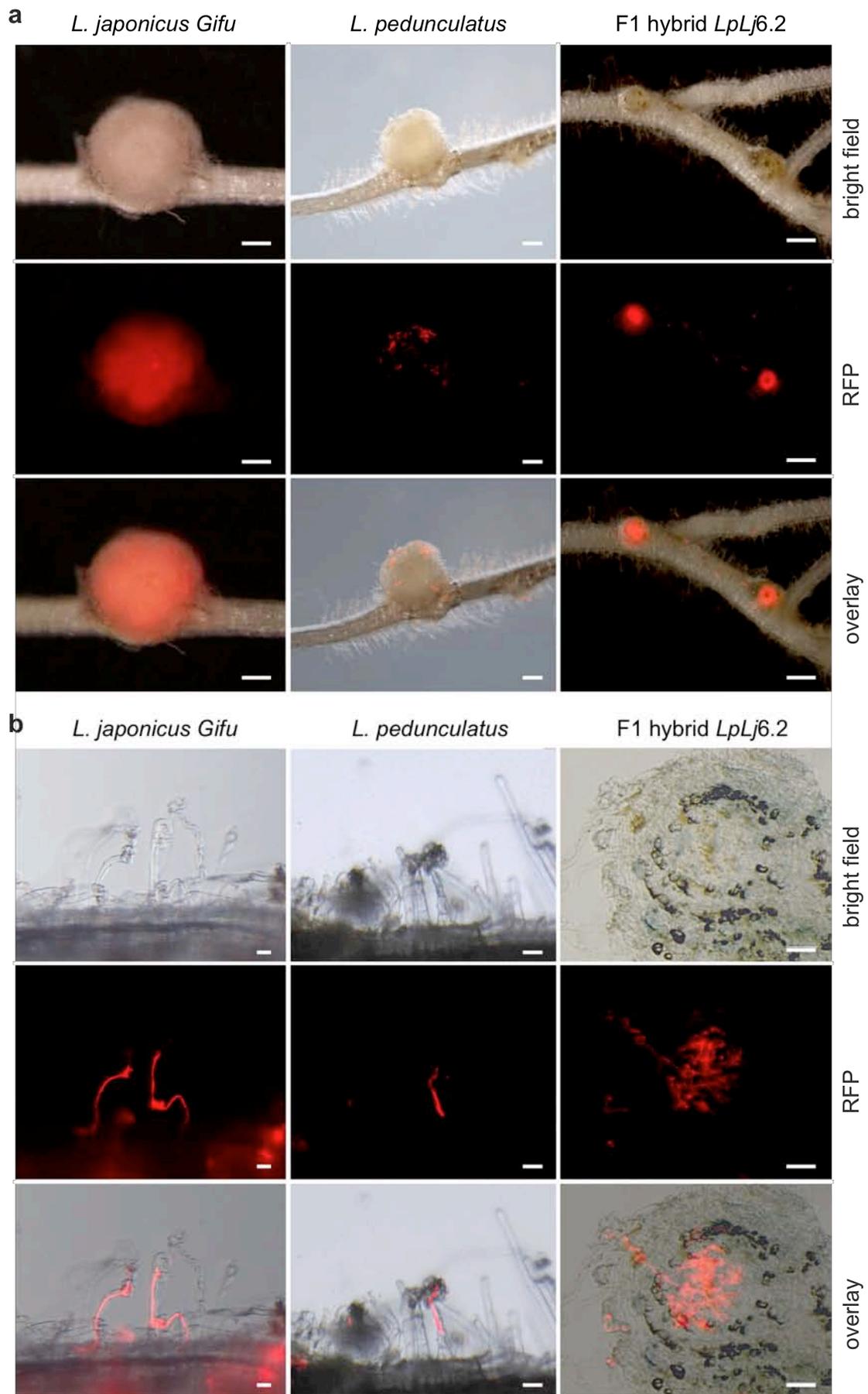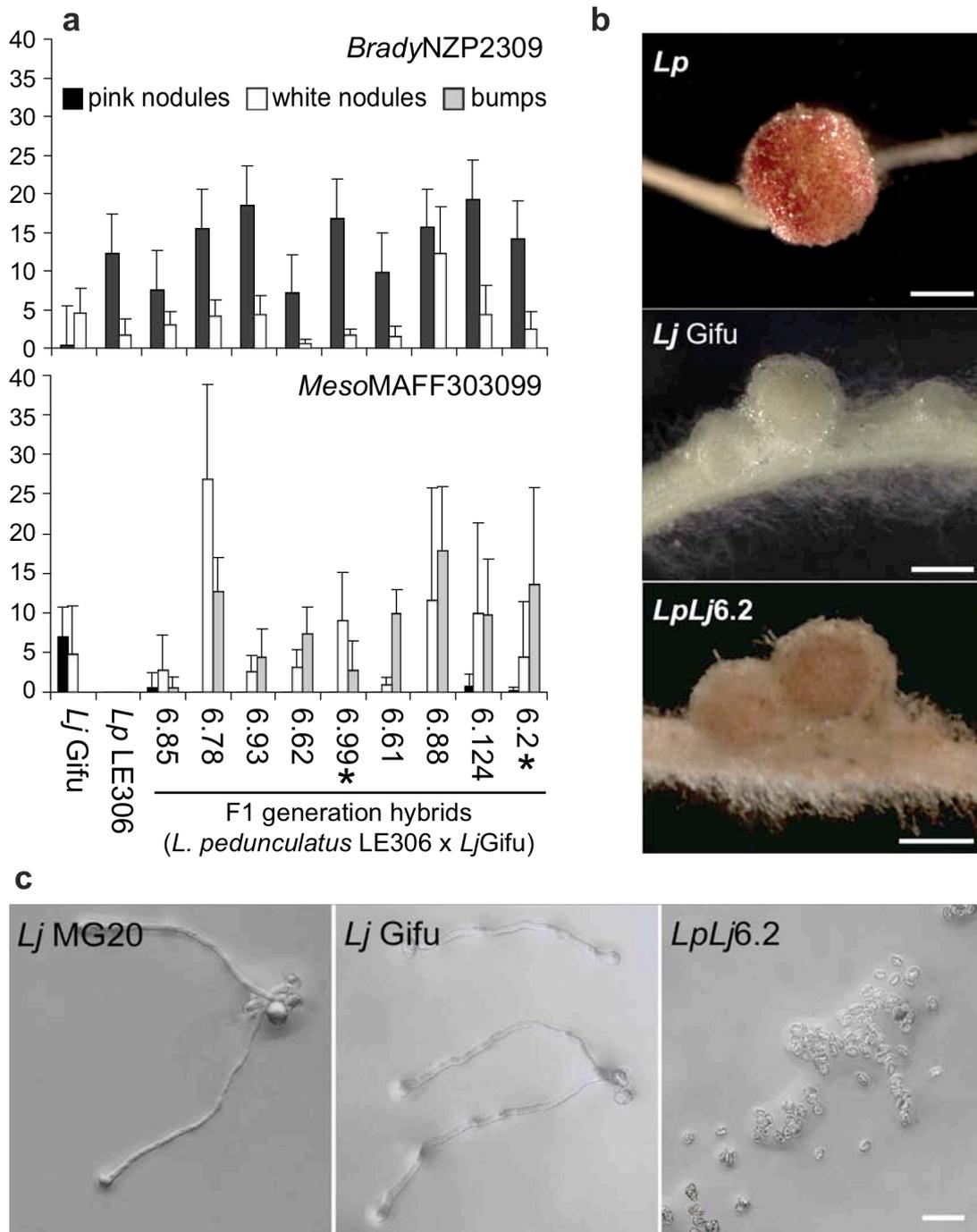
Figure 14. Nodulation and pollen grain phenotypes of *L. pedunculatus* LE306, *L. japonicus* Gifu and F1 generation hybrids.

(a), Quantitative nodulation phenotypes at 5 wpi. Five to ten seedlings per parental line and five to ten in-vitro-clones per F1 generation hybrid were inoculated with *M. loti* MAFF303099 (*Meso*MAFF303099) and *Bradyrhizobium* sp. strain NZP2309 (*Brady*NZP2309). The mean number of white and pink nodules and bumps was determined at 30 dpi. The F1 generation hybrids originate from a cross between *L. pedunculatus* (*Lp)* LE306 x *L. japonicus* (*Lj*) Gifu. Asterisk; hybrid with confirmed heterozygous genotype (III.1.3). (b), Root nodulation phenotypes induced by strain *Bradyrhizobium* sp. NZP2309 at 5wpi. *Lp* LE306, *L. pedunculatus* LE306; *Lj* Gifu, *L. japonicus* Gifu; *LpLj*6.2, F1 hybrid 6.2 from an original cross between *L. pedunculatus* LE306 x *L. japonicus; Lj* MG20*, L. japonicus* MG20. (c), Pollen grain germination test. Scale bars = 500 µm (b), 20 µm (c).

IV.2.2 Late infection blockage of *Meso*MAFF303099 by *LpLj*6.2 is mediated by the type III secretion system

M. loti effector proteins secreted via the type III secretion system have been shown to have a negative effect on nodulation in some legumes (Okazaki *et al.*, 2010). We investigated the nodulation phenotypes induced by DT3S, a *lacZ* labeled type III secretion system (T3SS) deletion mutant strain derived from *Meso*MAFF303099 (hereafter *Meso*DT3S). We inoculated in-vitro clones of *LpLj*6.2 and five other F1 hybrids with *lacZ* labeled *Meso*MAFF303099. In all F1 hybrids tested, infection by *Meso*MAFF303099 was blocked after the induction of organogenesis. Only bumps and small nodules formed along the root (Fig. 15). In contrast, *Meso*DT3S induced wild-type nodules on the F1 hybrids (Fig. 15), suggesting that the incompatibility between *Meso*MAFF303099 and the hybrid is due to recognition of molecules secreted by the T3SS. However, *Meso*DT3S failed to induce efficient nodules on *L. pedunculatus*, suggesting that additional molecules contribute to the incompatibility phenotype of *Meso*MAFF303099 on *L. pedunculatus* (Fig. 16a). On *L. japonicus* the nodulation efficiency levels of the wild-type strain *Meso*MAFF303099 and the mutant strain *Meso*DT3S are similar (Fig. 16a).

We performed co-inoculation experiments on *Lj* Gifu and *L. pedunculatus* with two mixtures of rhizobia: *Brady*NZP2309-*Meso*MAFF303099 and *Brady*NZP2309-*Meso*DT3S. Both *Meso*MAFF303099 and *Meso*DT3S alone induced a large number of white nodules on *L. pedunculatus* and equal numbers of white and pink nodules on *Lj* Gifu (Fig. 16a). *Brady*NZP2309, in contrast, induced a high number of white and only few pink nodules on *Lj* Gifu, and almost equal numbers of white and pink nodules on *L. pedunculatus*. Upon co-inoculation of *Lj* Gifu with *Brady*NZP2309 and *Meso*MAFF303099, the same number of pink nodules was induced as with the single strain inoculation of *Meso*MAFF303099, while more white nodules were induced. Upon co-inoculation of *Lj* Gifu with *Brady*NZP2309 and *Meso*DT3S, the number of white nodules was suppressed relative to the number of pink nodules, similar to *Meso*DT3S single strain inoculations.

Upon co-inoculation of *L. pedunculatus* with *Meso*MAFF303099 and *Brady*NZP2309, more white nodules were induced and slightly fewer pink nodules, leading to an increased difference in the number of pink and white nodules compared to single strain inoculation with *Brady*NZP2309. Upon co-inoculation of *L.*

pedunculatus with *Brady*NZP2309 and *Meso*DT3S, slightly more pink and white nodules were induced compared to single strain inoculation with *Brady*NZP2309. A potential inhibitory effect of nodulation on *L. pedunculatus* by *M. loti* T3SS is consistent with these observations. Although phenotypic differences are not statistically significant, the qualitative root infection phenotypes of *L. pedunculatus* differed greatly upon co-inoculation. The roots co-inoculated with *Brady*NZP2309-*Meso*DT3S formed nodules harboring *Meso*DT3S indicated by the blue staining for *lacZ* (Fig. 15). *Brady*NZP2309 co-inoculated with DsRed labeled *Meso*MAFF303099 induced root nodules devoid of *Meso*MAFF30309, demonstrated by aborted ITs in the nodule periphery and the absence of red fluorescence inside the nodule (Fig. 15).

In summary, we identified two levels at which *L. pedunculatus* controls the infection by *Meso*MAFF303099. The early infection incompatibility of *L. pedunculatus* to *Meso*MAFF303099 was rescued in the *LpLj*6.2 hybrid allowing the formation ITs and the IT progression to the cortex associated with the formation of primordia. However, the rhizobia were not released from the IT into the plant cell. This late infection blockage in the *LpLj*6.2 hybrid was abolished upon inoculation with a T3SS mutant strain *Meso*DT3S.
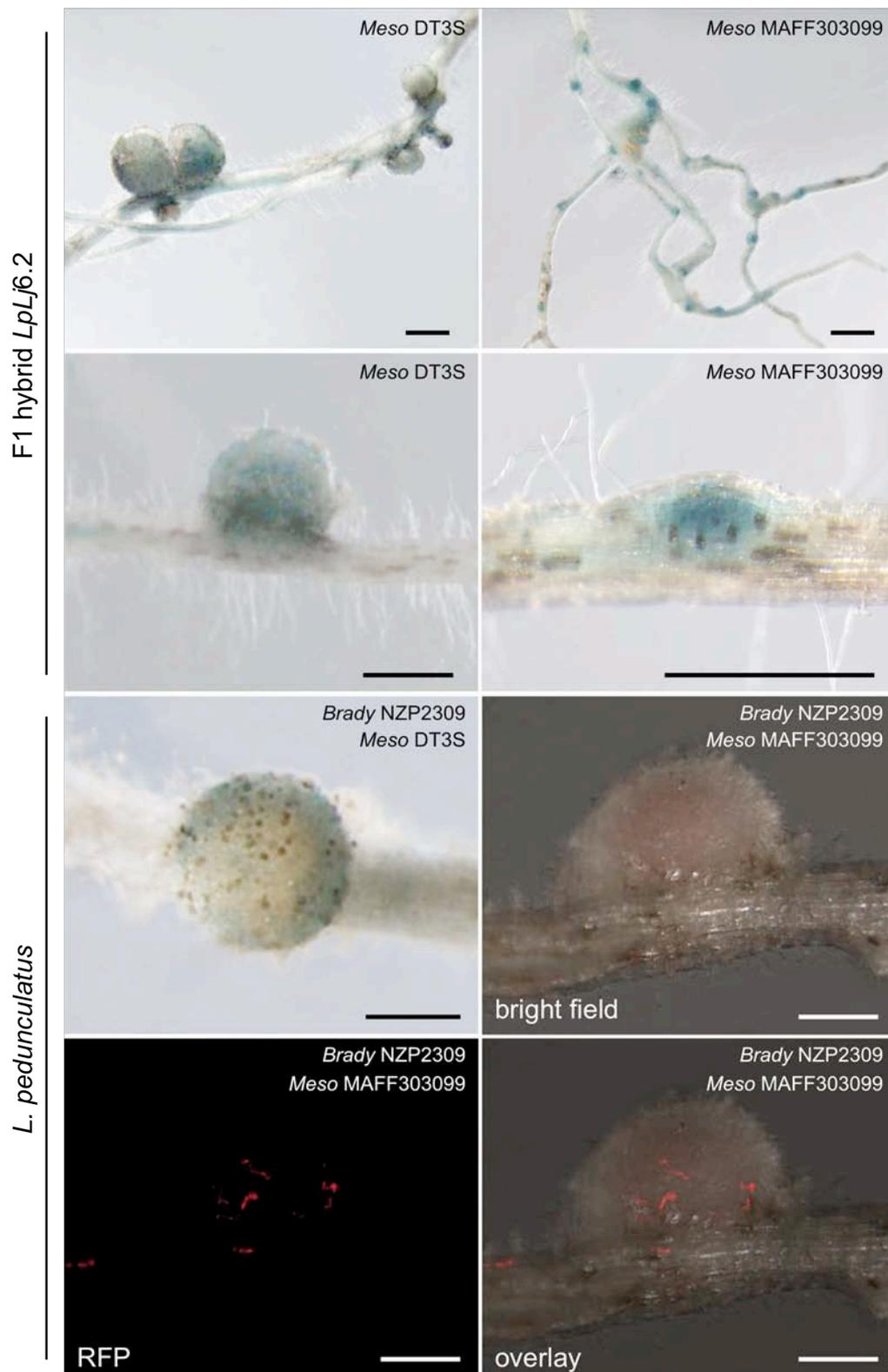
Figure 15. Root nodulation phenotypes at 5 weeks post single strain and co-inoculations.
Seedlings of *L. pedunculatus* LE306 and in-vitro clones of F1 hybrid 6.2 from an original cross between *L. pedunculatus* LE306 x *L. japonicus* Gifu were inoculated with single strains or co-inoculated with a mixture of rhizobial two strains. Blue X-Gal staining corresponds to of rhizobial lacZ expression. *Brady*NZP2309, *Bradyrhizobium* sp. NZP2309; *Meso*MAFF303099, DsRed-labeled *M. loti* strain MAFF303099; *Meso*DT3S, *lacZ*-labeled *M. loti* mutant strain DT3S; bright field, white light micrographs; RFP, red fluorescence micrographs; overlay, overlay of both white light and fluorescence micrographs visualizing infections with DsRed-labeled rhizobia. Scale bars = 500 μm.
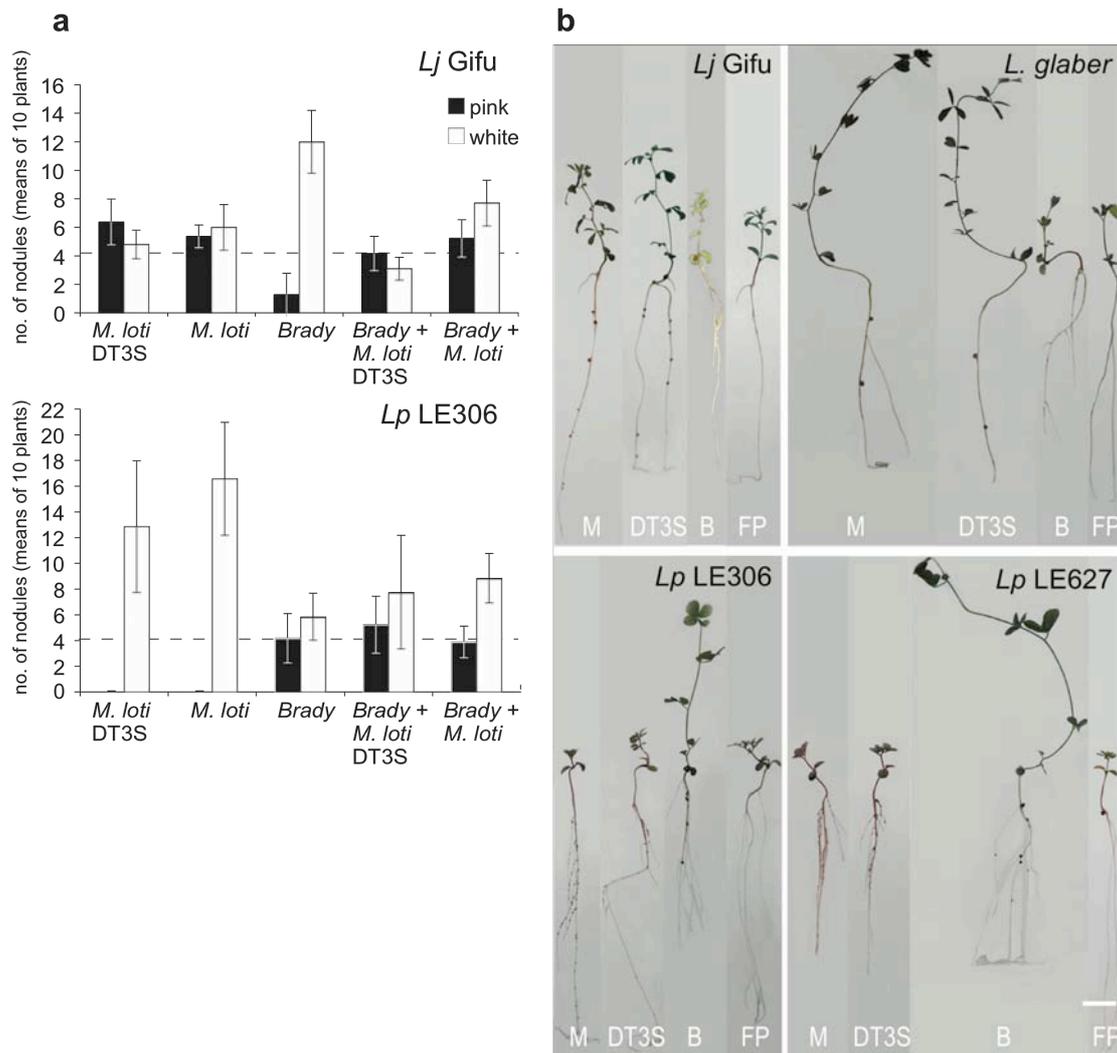
Figure 16. Quantitative nodulation phenotypes and whole plant phenotypes of *Lotus* spp. constituting contrasting compatibility groups.

(a), Quantitative nodulation phenotypes at 5 weeks post single strain and co-inoculations of *L. pedunculatus* LE306 and *L. japonicus* Gifu. Seedlings of *L. pedunculatus* LE306 and *L. japonicus* Gifu were inoculated with single strains or co-inoculated with a mixture of rhizobial two strains. *Brady*NZP2309, *Bradyrhizobium* sp. NZP2309; *Meso*MAFF303099, DsRed-labeled *M. loti* strain MAFF303099; *Meso*DT3S, *lacZ*-labeled *M. loti* mutant strain DT3S. Mean values are means of 10 plants. Error bars indicate standard deviations. (b), Plant phenotypes of *Lotus* species from different compatibility groups at 5 wpi. The plants were inoculated with *M. loti* strain, MAFF303099 (M), with the derived T3SS mutant strain (DT3S), and *Bradyrhizobium* sp. NZP2309 (B). *L. japonicus* Gifu (*Lj* Gifu) and *L. glaber* show efficient nodulation with both *M. loti* strains. The two *L. pedunculatus* cultivars (*Lp* LE306 and *Lp* LE627) form efficient symbioses with *Bradyrhizobium* sp. NZP2309. *Lp*, *L. pedunculatus*; *Lj* Gifu, *L. japonicus* Gifu. Scale bar = 1 cm.

IV.2.3 <u>Evolutionary history of *NFR5,* a candidate determinant of *Lotus* symbiont compatibility</u>

NFR5, in combination with *NFR1,* is a major RNS specificity determinant in *Lotus*. The ectodomain is generally assumed to be the region of NF binding in *NFR5*. We analyzed the nucleotide sequence variation of *NFR5* from a total of 30 individuals of natural European populations of *L. pedunculatus* and *L. corniculatus,* as well as five other closely related *Lotus* species (Table 9). Between one and four alleles were recovered per individual, consistent with the tetraploid status of some of the sampled *Lotus* species. In total, we recovered 61 *NFR5* alleles across the 30 sampled individuals.

Sequence polymorphism and interspecific divergence was evaluated across the *NFR5* coding region (Table 10; Fig. 18a). Among the 44 alleles isolated from *L. corniculatus*, the level of diversity at synonymous sites always exceeded that at non-synonymous sites. This is consistent with the action of purifying selection operating to remove deleterious mutations. The highest levels of synonymous variation (> 0.02) were found in the SP and LysM2 domains (for domain structure of NFR5 see Fig. 17-20), indicating that regions of this gene vary in their rate of evolution at synonymous sites. Interspecific divergence between *L. corniculatus* and *L. pedunculatus* also showed variation in the rate of synonymous evolution across domains. Nearly all domains showed greater divergence at synonymous sites than at non-synonymous sites, with the exception of the LysM1 and LysM2 domains. In these domains, divergence at non-synonymous sites exceeded that at synonymous sites (Table 10, Fig. 18a).

The phylogenetic reconstruction of the *NFR5* alleles from these species shows a clear separation of the alleles into two well-supported major clades according to host species origin (Fig. 19). Furthermore, these two clades correspond to different RNS compatibility groups: All species contributing alleles to the 'corniculatus' clade form efficient RNS with *M. loti,* while species contributing alleles to the 'pedunculatus' clade form efficient symbioses with *Bradyrhizobium* species.

Table 9: Genbank accession numbers for *Lotus* spp. *NFR5* and the orthologs *MtNFP* and *PsSYM10*

Plant species	Clone ID[a/b]	Accession No.
Medicago sativa		DQ496250
Pisum sativum cv. Alaska		AJ575250
Pisum sativum cv. Frisson		AJ575251
Pisum sativum cv. Sparkle		AJ575252
Pisum sativum cv. Finale		AJ575253
Lotus japonicus ecotype Gifu		AJ575254
Lotus burttii B-303		HQ448883[c]
Lotus filicaulis		HQ448884[c]
Lotus glaber cv. Herminia	K729_10B[a]	HQ448885[c]
Lotus glaber cv. Herminia	K729_9B	HQ448886[c]
Lotus glaber cv. Herminia	K725_1A	HQ448887[c]
Lotus japonicus ecotype Gifu		HQ448888[c]
Lotus japonicus ecotype MG20		HQ448889[c]
Lotus japonicus ecotype NepalII		HQ448890[c]
Lotus subbiflorus cv. El Rincon	K712_18C	HQ448891[c]
Lotus subbiflorus cv. El Rincon	K715_28D	HQ448892[c]
Lotus pedunculatus cv. Maku	K734_37E	HQ448893[c]
Lotus pedunculatus cv. Maku	K743_33E	HQ448894[c]
Lotus pedunculatus cv. Maku	K743_47F	HQ448895[c]
Lotus pedunculatus	S5P1c15[b]	HQ448896[c]
Lotus pedunculatus	S5P1c28	HQ448897[c]
Lotus corniculatus	S1P5c25	HQ448898[c]
Lotus corniculatus	S1P5c28	HQ448899[c]
Lotus corniculatus	S1P5c46	HQ448900[c]
Lotus corniculatus	S2P1c22	HQ448901[c]
Lotus corniculatus	S2P1c23	HQ448902[c]
Lotus corniculatus	S2P5c1	HQ448903[c]
Lotus corniculatus subsp. *alpinus*	S4P1c1	HQ448904[c]
Lotus corniculatus subsp. *alpinus*	S4P1c7	HQ448905[c]
Lotus corniculatus	S6P1c33	HQ448906[c]
Lotus corniculatus	S7P1c1	HQ448907[c]
Lotus corniculatus	S7P2c23	HQ448908[c]
Lotus corniculatus	S7P2c25	HQ448909[c]
Lotus corniculatus	S9P1c22	HQ448910[c]
Lotus corniculatus	S9P1c23	HQ448911[c]
Lotus corniculatus	S9P2c5	HQ448912[c]
Lotus corniculatus	S9P2c7	HQ448913[c]
Lotus corniculatus	S9P4c16	HQ448914[c]
Lotus corniculatus	S9P4c4	HQ448915[c]
Lotus corniculatus	S12P1c1	HQ448916[c]
Lotus corniculatus	S12P1c17	HQ448917[c]
Lotus corniculatus	S12P1c8	HQ448918[c]
Lotus corniculatus	S13P1c48	HQ448919[c]
Lotus corniculatus	S13P2c3	HQ448920[c]
Lotus corniculatus	S13P5c2	HQ448921[c]
Lotus corniculatus	S13P5c54	HQ448922[c]
Lotus corniculatus	S13P5c55	HQ448923[c]
Lotus corniculatus	S16P1c14	HQ448924[c]
Lotus corniculatus	S16P1c15	HQ448925[c]
Lotus corniculatus	S16P2c51	HQ448926[c]
Lotus corniculatus	S16P2c53	HQ448927[c]
Lotus corniculatus	S16P2c6	HQ448928[c]
Lotus corniculatus	S17P2c2	HQ448929[c]

Clone nomenclature: [a] Plant ID_clone number and [b] S: site of collection, [c] sequences generated during this study ,P: plant number, c: clone number

Table 10: Nucleotide polymorphisms – DnaSP v5 (Librado & Rozas, 2009)

clades tested	alleles	sites without gaps	haplo-types	Pi total (sites)	Pi non (sites)	Pi syn (sites)	
corniculatus	44	1764	39	0,0047 (63)	0,00278 (31)	0,01082 (32)	coding region
Clade 1	53	1756	43	0,00454 (65)	0,00283 (33)	0,01001 (32)	
Clade 2	8	1764	6	0,0216 (91)	0,01263 (53)	0,05030 (41)	
pedunculatus	5	1764	4	0,00204 (6)	0,00134 (3)	0,00429 (3)	
corniculatus	44	78	6	0,0088 (5)	0,00382 (3)	0,02254 (2)	signal peptide (1-78)
Clade 1	53	78	6	0,00746 (5)	0,00319 (3)	0,01926 (2)	
Clade 2	8	78	3	0,02473 (4)	0,02599 (1)	0,02125 (3)	
pedunculatus	5	78	1	0 (0)	0 (0)	0 (0)	
corniculatus	44	141	5	0,0016 (4)	0,00127 (2)	0,00255 (2)	LysM1 (154-294)
Clade 1	53	141	5	0,00133 (4)	0,00106 (2)	0,00212 (2)	
Clade 2	8	141	2	0,00532 (3)	0,00236 (2)	0,01422)1)	
pedunculatus	5	141	1	0 (0)	0 (0)	0 (0)	
corniculatus	44	141	6	0,0095 (6)	0,00439 (2)	0,02618 (4)	LysM2 (340-480)
Clade 1	53	141	6	0,01094 (6)	0,00671 (2)	0,02476 (4)	
Clade 2	8	141	5	0,03875 (13)	0,02583 (8)	0,08071 (7)	
pedunculatus	5	141	3	0,00851 (2)	0,00556 (1)	0,01812 (1)	
corniculatus	44	132	6	0,00239 (6)	0,00136 (3)	0,00562 (3)	LysM3 (541-672)
Clade 1	53	132	6	0,00199 (6)	0,00113 (3)	0,00468 (3)	
Clade 2	8	132	3	0,02435 (7)	0,00430 (6)	0,08616 (1)	
pedunculatus	5	132	1	0 (0)	0 (0)	0 (0)	
corniculatus	44	519	15	0,00443 (19)	0,00234 (8)	0,01095 (11)	All LysM (154-672)
Clade 1	53	519	16	0,00458 (19)	0,0028 (8)	0,00993 (11)	
Clade 2	8	519	5	0,02154 (27)	0,01051 (18)	0,05631 (11)	
pedunculatus	5	519	3	0,00231 (2)	0,00152 (1)	0,00480 (1)	
corniculatus	44	75	4	0,00678 (3)	0,00640 (2)	0,00754 (1)	TM (739-813)
Clade 1	53	75	4	0,00579 (3)	0,00553 (2)	0,00630 (1)	
Clade 2	8	75	3	0,04238 (8)	0,03430 (3)	0,06067 (5)	
pedunculatus	5	75	1	0 (0)	0 (0)	0 (0)	
corniculatus	44	948	21	0,00353 (28)	0,00157 (17)	0,01014 (11)	Kinase (814-1788)
Clade 1	53	946	24	0,0032 (30)	0,00144 (18)	0,00922 (12)	
Clade 2	8	948	5	0,02132 (48)	0,01207 (29)	0,05274 (20)	
pedunculatus	5	951	3	0,00252 (4)	0,00163 (2)	0,00555 (2)	

Note. Pi: nucleotide diversity, non: non-synonymous polymorphisms, syn: synonymous polymorphisms; Clade 1 comprises *L. corniculatus, L. glaber, L. japonicus, L. burttii, L. filicaulis* alleles; Clade 2 comprises *L. pedunculatus a*nd *L. subbifloru*s alleles.

```
           10        20        30        40        50        60        70        80        90       100
MAVFFLTSGSLSLFLALTLLFANIAARSEQISGTDFSCPVDSPPSCETYVSYTAQSPNFLSLTNISDIFDISPLSIARASNIDAGKDKLVPGQVLLVPVT
......P...Q........-.T...................-....T............................................................
.........-.T...................-....T.......................................................................
.........-.T...................-....T.......L...............................................................
.........-.T......K...P........-....T.......L...............................................................
.........-.T......K...P........-....T.......L...............................................................
.........T.....K...P.............T.......L.................................................................
                                    hom
          110       120       130       140       150       160       170       180       190       200
CGCTGNHSFANTSYQIQLGDSYNFIATTSYENLTNWNIVQDSNPGVNPYLLPEGIKVVFPLFCRCPSKNQLNKGIQYLITYVWKPNDNVSLVSAKFGASP
.................D.V...L..................................................................................
.................D.V...L..................................................................................
...A....S.......K...D.V...L.....A...............RV.........................................................
...A....S.......K...D.V...L.....A...............RV.........................................................
...A....S.......D.V...L.....A...............RV.........................................................
...A....S.......D.V...L.....A...............RV.........................................................
          210       220       230       240       250       260       270       280       290       300
ADILTENRYGQDFTAATNLPVLIPVTQLPELTQPSSNGRKSSIHLLVILGITLGCTLLTAVLTGTLVYVYCRKKKALSRTASSAETADKLLSGVSGYVSK
.......--.................................................S.....R....N...................................
.......--...................................................R....N.......................................
..........I.................................................R....N.......................................
.......--....I..............................................R....N.......................................
.......--....I..............................................R....N.......................................
............I...............................................R....N.......................................
          310       320       330       340       350       360       370       380       390       400
PNVYDIDEIMEATKNFSDECKVGESVYKANIEGRVVAVKKIKEGGANEELKILQKVNHGNLVKLMGVSSGNDGNCFLVYEYAENGSLAEWLFSKSSGTSN
...E.........D.....................-......................................................................
...E.........D.....................-......................................................................
...E.........D..............................................Y..........................................P.
...E.........D.....................-........................Y..........................................P.
...E.........D.....................-........................Y..........................................P.
...E.........D..............................................Y..........................................P.
                                    hom
          410       420       430       440       450       460       470       480       490       500
                                                             *
SLTWSQRISIAVDVAVGLQYMHEHTYPRIIHRDITASNILLDSTFKAKIANFAMARTSTNPMMPKIDVFAFGVLLIELLTGRKAMTTKENGEVVMLWKDM
...............................T.........................................................................
...............................T.........................................................................
...............................T.........................................................................
...............................T.........................................................................
...............................T.........................................................................
...............................T......N.................................................................
          510       520       530       540       550       560       570       580       590   595
WEIFDIEENREERIKKWMDPQLESFYHIDNALSLASLAVNCTADKSLSRPSMAEIVLSLSFLSQQSSNPTLERSLTSSGLDVEEDAQTVTSITAR   Lp
.............K..................................T.........T.P.-.......-.......---HI-------   node A
.............K..................................T...-.......-.......---HI-------   node B
.............K..................................T...-.......-.......---HI-------   node C
.............R.....N............................T...-.......-.......---HI-------   node D
.............R.....N............................T...-.......-.......---HI-------   node E
.............R.....N............................T...................D..HI.......   Lj
```

Figure 17. Multiple alignment of reconstructed ancestral *Lotus* sequences to *L. pedunculatus* and *L. japonicus* Gifu.

The ancestral *Lotus* sequences were calculated at the ancestral nodes A-E of the phylogenetic tree (Fig. 19). Missing characters were not calculated for the reconstructed sequences due to gaps in original alignment. Shaded sections of the alignment correspond to the NFR5 domain structure. SP, signal peptide; LysM1 to LysM3, lysine motif domains one to three; TM, transmembrane domain; KD, kinase domain; (.), character matching the sequence of the first row; (-), missing character; Lp, *L. pedunculatus* clone K0743_47F; Lj, *L. japonicus* Gifu.
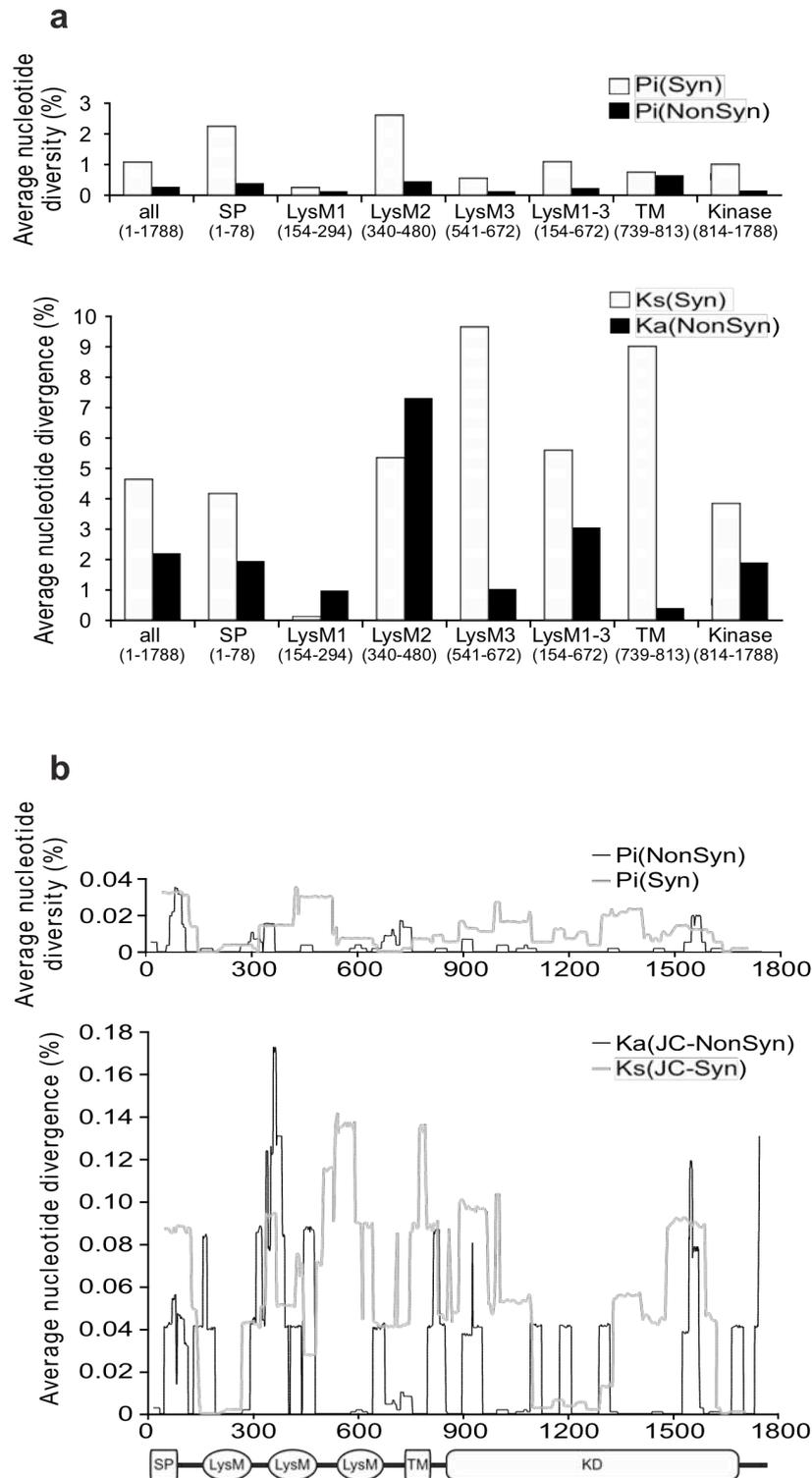
Figure 18. Genomic diversity and divergence at the *NFR5* locus.
(a), average nucleotide diversity (Pi) between *L. corniculatus* alleles and nucleotide divergence (K) of *L. pedunculatus* to *L. corniculatus* was individually calculated from 61 alleles of *Lotus* species for each *NFR5* domain. (b) Sliding window analysis of overall sites of *NFR5*. Nucleotide polymorphisms of 61 alleles of six *Lotus* species were analyzed with a window length of 25 and step size 1. Abscissa indicates nucleotide position. Ordinate indicates nucleotide diversity (Pi) between *L. corniculatus* alleles and nucleotide divergence (K) of *L. pedunculatus* to *L. corniculatus.* JC, Jukes and Cantor correction; Non-syn, non-synonymous nucleotide polymorphisms (resulting in amino acid substitutions); Syn, synonymous nucleotide polymorphisms (without amino acid substitution); SP, signal peptide; LysM1 to LysM3, lysine motif domains one to three; TM, transmembrane domain; KD, kinase domain.
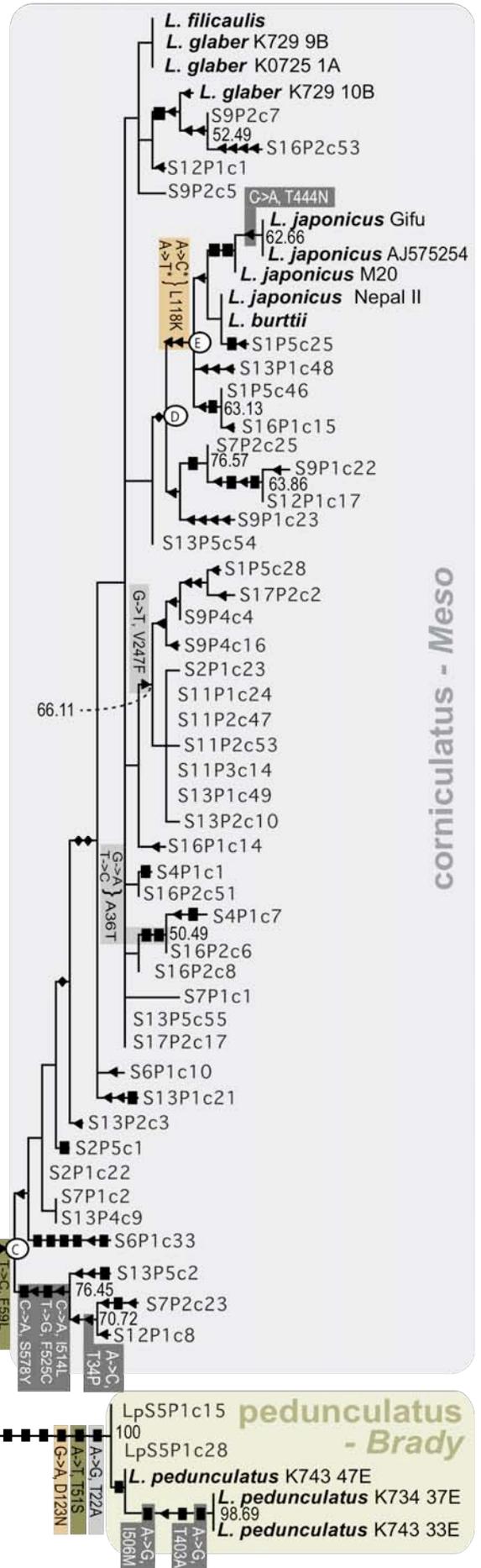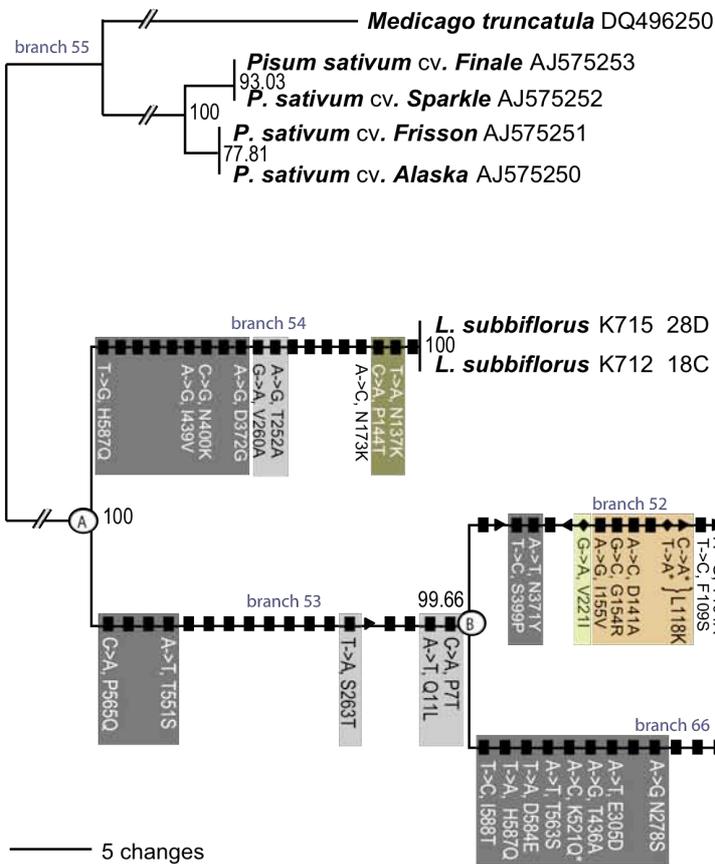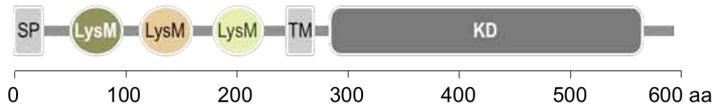
Figure 19. *NFR5* gene phylogeny.

The phylogenetic tree was calculated from a multiple alignment of 61 alleles of European *L. corniculatus* and *L. pedunculatus* populations, and reference sequences. The tree was constructed under the maximum parsimony (MP) criterion and represents one of 100 most parsimonious trees. NFR5 orthologs from *M. truncatula* (*NFP*) and *P. sativum* (*SYM10*) were used as outgroup. Bootstrap probability (BP) values below 60 are not displayed. Using ancestral state reconstruction methods unambiguously changing characters (nucleotide states) were traced in the *NFR5* phylogenetic tree. Unambiguous character changes are mapped to the respective branch. Labels for unambiguous changes resulting in amino acid substitution (non-synonymous nucleotide substitutions in the multiple alignment) are superimposed onto the respective change. Shading colors correspond to NFR5 protein domains as shown in inset. Ancestral sequences were calculated for nodes with encircled capital letters A-E. SP, signal peptide; LysM, lysine motif domain; TM, transmembrane domain; KD, kinase domain.

We used ancestral state reconstruction to determine whether particular substitutions at this gene were associated with differentiation of symbiont recognition. Unambiguous nucleotide changes were mapped onto the phylogenetic tree (Fig. 19). Amino acid changes are indicated by a shaded background corresponding to the color of the NFR5 protein domains in which the substitution occurred. A large proportion of the amino acid changes along branch 52, which subtends the 'corniculatus' clade, occur in the LysM2 domain. Changes along this branch may be associated with the discrimination between mesorhizobial and bradyrhizobial symbionts. Of particular interest is the amino acid substitution K118L along the branch subtending the 'corniculatus' clade. This amino acid position was identified in a study testing the compatibility phenotype of DZL strain on *L. japonicus* and *L. filicaulis* (Radutoiu *et al.*, 2007). Following a series of domain swaps and site-directed mutagenesis studies, the authors concluded that the leucine/lysine difference at position 118 was largely responsible for the differential response of *L. japonicus* and *L. filicaulis* towards NF produced by DZL (Radutoiu *et al.*, 2007). This amino acid change along branch 52 is caused by two nucleotide substitutions at a single codon position. These two changes are homoplastic: the clade containing the alleles from *L. japonicus, L. burttii* and four *L. corniculatus* alleles (node D, Fig. 19) shows a reversal to the ancestral state at this position, while *L. filicaulis* expresses the derived state.

The ratio of non-synonymous/synonymous substitution rates ($\omega = d_N/d_S$) is a widely used measure of selective pressure at the nucleotide level (Nei & Kumar, 2000). A value of $\omega > 1$ is indicative of positive Darwinian selection, while $\omega < 1$ is indicative of purifying selection. Positive selection can result in the adaptive fixation of amino acid substitutions between species, accelerating the rate of evolution of non-synonymous changes relative to synonymous changes. Based on our observations of the average interspecific Ka/Ks ratios (similar to ω) across domains (Fig. 18a) and the

sliding window analysis (Fig. 18b), we identified the LysM2 domain and the inter-domain region of LysM1 and LysM2 as hot spots for evolution at non-synonymous sites. To further evaluate the statistical significance of this signature of positive selection, we conducted a series of codon-based maximum likelihood analyses implemented in PAML (Nielsen & Yang, 1998; Yang *et al.*, 2000).

To test whether the rate of substitution and/or the ω value (the non-synonymous/synonymous rate ratio) differed between the LysM domains and the rest of the gene, we applied the "fixed-sites" models A-E described in (Yang & Swanson, 2002). We partitioned the dataset into two sets: one set including the regions encoding LysM domains/LysM2 domain and a second set including the regions not encoding LysM domains/LysM2 domain. Yang and Swanson (2002) defined six models, which can be compared to determine whether sequence evolution differs significantly between gene partitions. For example, by comparing model A and B, it is possible to determine if the two partitions differ in their overall rate of substitution, by assuming proportional branch lengths among partitions (Yang and Swanson, 2002). In our analysis of the *NFR5* gene, the partition containing LysM domains has a rate of 2.74 higher than the rest of the gene (Table 11). However, the log-likelihood value was not significantly larger for this model compared to the null model (Model A). The use of particular pairwise comparisons among the other six models allows one to evaluate if and how other aspects of sequence evolution differ between partitions, for example if the transition transversion ratio, the non-synonymous/synonymous rate ratio or codon frequencies differ between partitions. Since none of the models with these additional parameters provided a better fit to our data, we failed to reject the null hypothesis that the two gene partitions (regions containing LysM domains versus those that do not) differed significantly in their sequence evolution (Table 11).

Table 11: Log-likelihood values and parameter estimates for the *NFR5* gene under fixed-sites models

Model	p	lnL	branch length partition 1	branch length partition 2	r_2	k (ts/tv)	ω	
A (homogeneous)	100	-3705.02	0.33	0.33	1.00	2.54	0.35	
B (different *r*s)	101	-3704.90	0.34	0.92	2.74	2.53	0.35	
C (different *r*s and πs)	110	-3701.96	0.33	0.94	2.81	2.54	0.35	
D (different rs, k and πs)	103	-3704.55	0.34	0.92	2.74	2.38	0.33	non-LysM
						3.06	0.41	LysM
E (different rs, k, ω, and πs)	112	-3701.57	0.33	0.94	2.81	2.38	0.33	non-LysM
						3.10	0.41	LysM

Note. p: Number of parameters including branch lengths b = 89 in the tree. The two partitions are the non-LysM sites and the LysM sites. r_2 is the rate of the second site partition relative to the rate of the first partition.

PAML analyses can also be used to identify targets of positive selection, in the absence of detailed information about putative selective targets. The model M7 described in Yang and Swanson (2002) assumes a beta distribution of ω values across all sites, in which ω is bounded by 0 and 1. Model M8 allows for a beta distribution of ω values and an extra site class, in which ω is not bounded between 0 and 1. Model M8 provided a statistically significant better fit to the *NFR5* data (2Δl = 2 x ((-3677.58) - (-3686.56)) = 17.95, d.f. = 2, p = 0.0001), implying that some sites within the *NFR5* gene have experienced positive selection. Nearly 5% of the sites within *NFR5* had ω>1 and most of these were located in the LysM2 domain and the kinase domain (Table 12, Fig. 20). Two methods, the naïve empirical Bayes (NEB) and the Bayes empirical Bayes (BEB), can be used to calculate the posterior probabilities of putatively selected sites. In our analyses two sites (K30 and N521) showed NEB and BEB values greater than 0.95 and are likely to be under positive selection. An additional 26 sites have posterior probabilities 0.50<BEB<0.95 and NEB<0.5, of which P34 and E305 have ω>2. Sites with posterior probabilities close to zero and posterior means ω<0.1 are likely to be under purifying selection, in which mutations at these sites are deleterious and are removed by natural selection.

Extended models, called 'branch-site models' allow the ω ratios to vary among lineages and among sites. They can be used to identify positive selection along pre-specified lineages; so-called 'foreground-branches' (Yang & Nielsen, 2002). We specified the branch 52 (Fig. 19) as the 'foreground-branch' (the branch for which to test positive selection) and identified six positively selected sites (Table 13). Three of these sites (59L, 118L, 399P) were also identified using random-sites models, but with varying support values (Table 12). Interestingly, in a separate analysis we tested for positive selection along the *Medicago-Pisum* branch and identified positive selected sites different than the ones in branch 52 but residing in close vicinity, supporting that these protein regions are under positive selection and that the different sites residing within this region explore different selective forces depending on the rhizobial symbiont.

Table 12: Log-likelihood values under two random-sites models for the *NFR5* gene

Model	p	lnL	Positively Selected Sites with P(ω>1)	
			NEB	BEB
M7 (beta)	101	- 3686.56	Not allowed	
M8 (beta & ω>1)	103	- 3677.58	**30K, 521N**	*6L,* **30K,** *34P, 59L, 97V, 104A, 106N, 118L, 125V, 221I, 230E, 305E, 339K[a], 515R,* **521N**

Note. p: Number of parameters including branch lengths b = 89 in the tree. NEB: Naive Empirical Bayes, BEB: Bayes Empirical Bayes. Sites inferred to be under positive selection at 95% level (bold) and below 95% (italic). Reference sequence is that of *L. japonicus* Gifu (see Fig. 17). [a] only two *L. corniculatus* alleles (S12P1c1 and S12P1c8) carry a methionine at position 339.
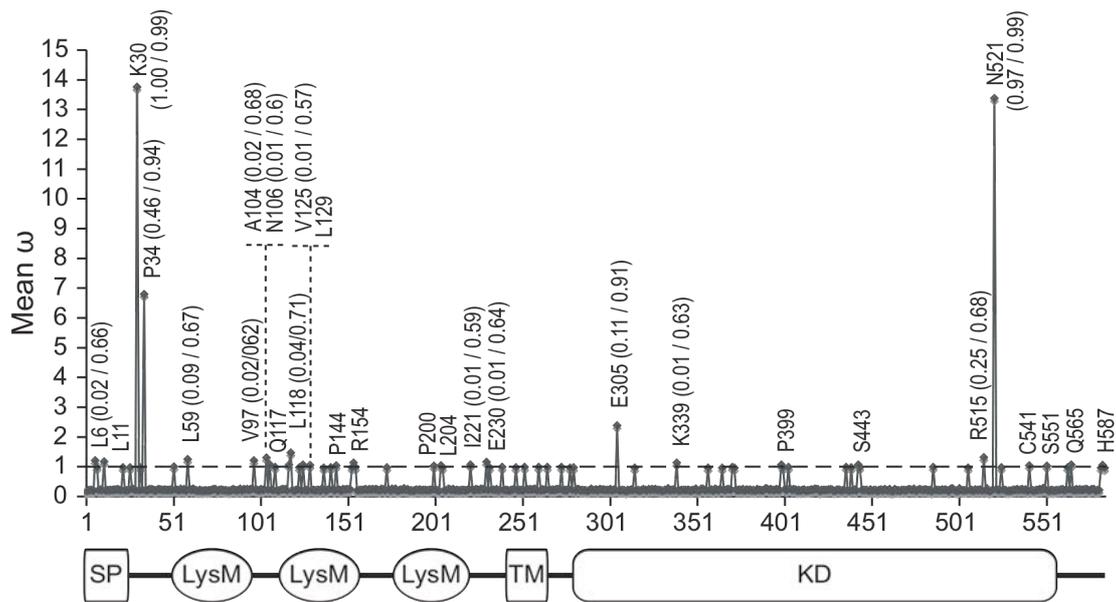
Figure 20. Distribution of the posterior means of ω for sites (codons) along the *NFR5* gene.
The posterior means were calculated under the random-site model M8 (beta & ω) (Yang *et al.*, 2000). The ordinate represents the average of ω over 11 site classes, weighted by posterior probabilities (Yang & Swanson, 2002). For sites with ω>1, amino acids corresponding to the *Lj* Gifu sequence and Naive and Bayes Empirical Bayes (NEB / BEB) probability values >0.5 are superimposed on the graph. The protein domain structure is displayed below the graph. SP, signal peptide; LysM1 to LysM3, lysine motif domains one to three; TM, transmembrane domain; KD, kinase domain.

Table 13: Parameter estimates for branch-site model

Foreground branch	Branch-Site Model A	lnL	Branch-Site Test of Positive Selection (LRT2)	Positively Selected Sites for Foreground Lineages $P(\omega>1)$	Bayes Empirical Bayes (BEB) probability	
branch 52	null hypothesis ($\omega_2=1$ fixed)	- 5558.82		57 (59) L[a]	0.50	LsyM1
				107 (109) S	0.63	
			6.51 (P=0.039)	116 (118) K	0.59	LsyM2
	alternative hypothesis ($\omega_2 \geq 1$)			139 (141) A	0.53	LsyM2
				366 (371) Y	0.64	KD
		- 5555.56		394 (399) P	0.61	KD
branch 55	null hypothesis ($\omega_2=1$ fixed)	- 5554.67		12 A[b] (12) S	0.64	SP
				28 L (28) E	0.66	
			5.31 (P=0.07)	84 D (86) K	0.52	LsyM1
				85 K (87) D	0.54	
	alternative hypothesis ($\omega_2 \geq 1$)			134 Y (136) W	0.61	LsyM2
		- 5552.01		136 E (138) I	0.61	LsyM2
				143 N (145) G	0.76	LsyM2
				260 L (264) G	0.72	KD
				272 R (276) A	0.78	KD
				315 N (319) E	0.53	KD
				342 S (347) N	0.70	KD
				392 K (397) G	0.57	KD
				495 F (500) M	0.61	KD
				562 E (568) N	0.53	KD

Note. [a] amino acid sequence referring to node C, the hypothetical ancestral state of the 'corniculatus' clade; [b] amino acid sequence referring to node A, the hypothetical ancestral state of *Lotus* clade. Brackets: Reference sequence is that of *L. japonicus* Gifu (see Fig. 17).

IV.2.4 Polymorphic sites reside in surface exposed regions of NFR5

Several lines of evidence indicate that the LysM2 domain is a target of positive selection. We used homology modeling to examine the structural context of the potential adaptive substitutions (Fig. 21). Non-synonymous changes distinguishing *NFR5* alleles of *L. corniculatus* and *L. pedunculatus* are color-coded. These substitutions correspond to those inferred by ancestral state reconstruction to have evolved along the "corniculatus" and "pedunculatus" lineages, respectively. All substitutions within the LysM2 domain were predicted to be surface exposed and distributed on both faces of the molecule. The LysM2 domains of *M. truncatula* and *P. sativum* are only 45.1% sequence identical to those of *Lotus* (23 identical amino acids, 28 amino acid substitutions in either *M. truncatula* or *P. sativum* or in both). The positively selected sites along branch 55 (Fig.19) were identified under various models (Table 12; Table13; Fig. 19; Fig. 20) and were predicted to be surface

exposed and to reside in close vicinity to positively selected sites identified for the "corniculatus" and "pedunculatus" lineages (Fig. 22).

High non-synonymous substitution rates and sites under positive selection were also found within the kinase domain (Fig. 18b; Fig. 20). We used homology modeling to examine the structural context of the potential adaptive substitutions within the kinase domain (Fig. 23). From a total of 13 sites, which distinguished the *NFR5* alleles *Lj* Gifu and *L. pedunculatus* in the kinase domain, nine could be placed onto the homology model. The remaining four sites were outside the region available for threading. Of the 9 positions that could be placed on the model, five sites had signatures of positive selection (Fig. 23, red lettering). Most of the amino acid substitutions were predicted to reside in surface exposed loops. The sites under positive selection were predicted to cluster together in two regions of the kinase (Fig. 23, dashed lines).
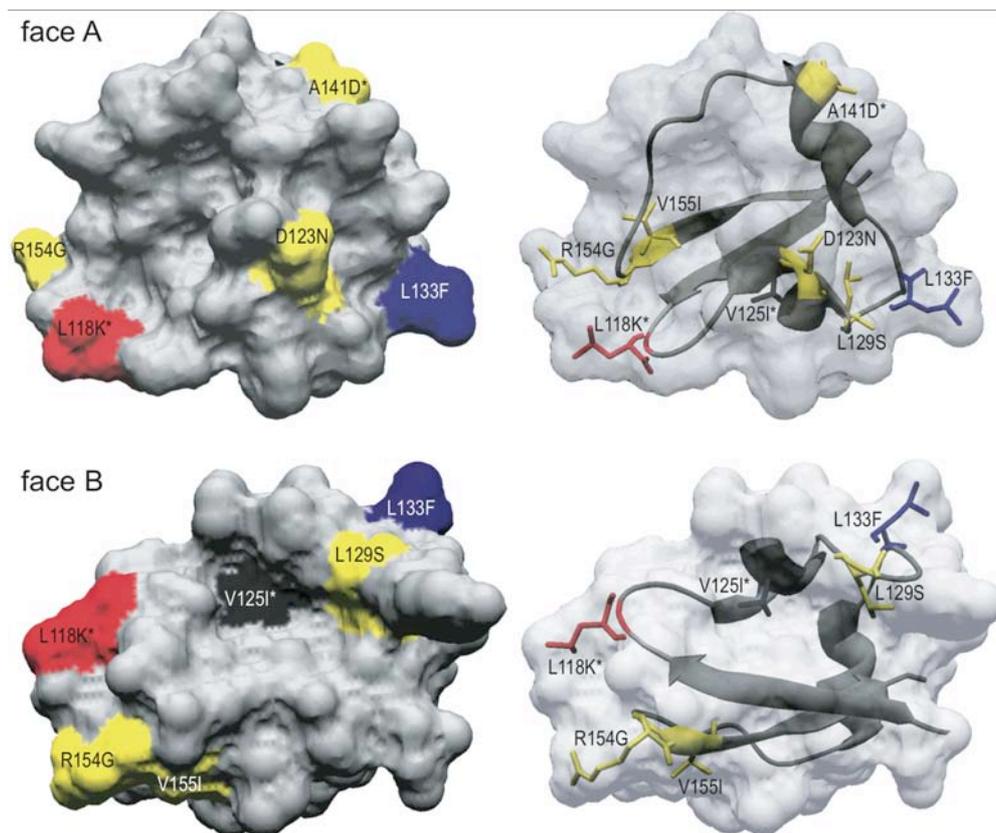
Figure 21. Homology model of LysM2 domain of *Lj* Gifu-NFR5.
The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD, pdp ID:1e0gA) served as template. Red and yellow, amino acid substitutions among 'corniculatus' clade and 'pedunculatus' clade (Fig. 19); red, amino acid substitution with reversion (A->C and A->T => L118K) at ancestral node. This amino acid substitution was implicated in determining rhizobium specificity in *L. filicaulis* (Radutoiu *et al.*, 2007). Black, the amino acid polymorphism (V125I) differing within the 'pedunculatus' clade; blue, gain-of-function mutation at homologous site in *At-CERK1* that leads to a deregulated defense response upon barley powdery mildew treatment (Lipka, personal communication, October 16, 2008); asterisk, amino acid substitution at site under positive selection.
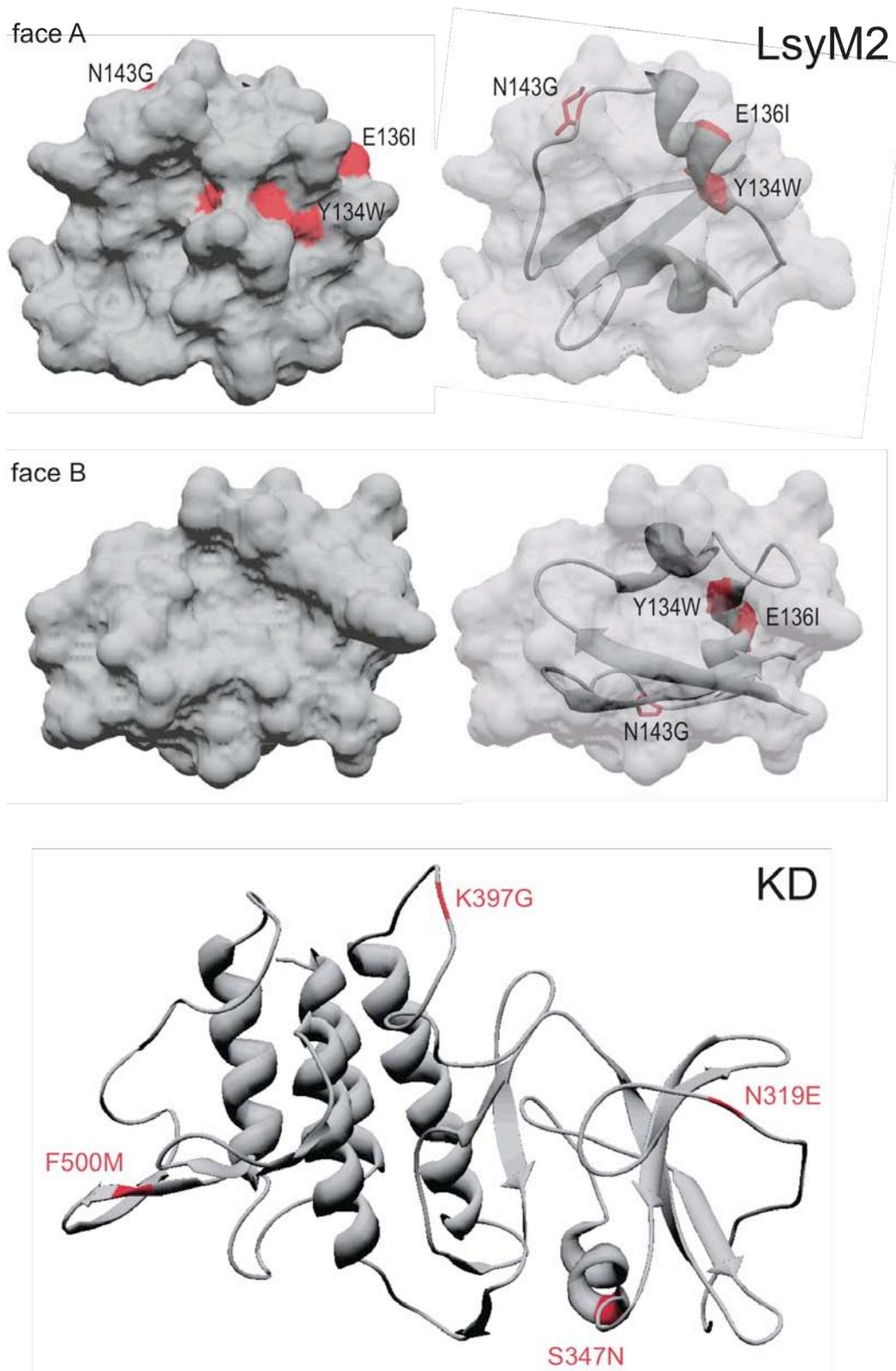
Figure 22. Homology models of LysM2 and the kinase domain of *Lj* Gifu-NFR5.
The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD, pdp ID:1e0gA) served as template for the homology model of *Lj* Gifu-NFR5 LysM2 domain. The crystal structure of tomato Pto (PDB ID: pdb3hgk) (Dong *et al.*, 2009) served as template for the homology model of *Lj* Gifu-NFR5 kinase domain. Red, amino acid substitutions between *P. sativum-M. truncatula* clade and the *Lotus* clade with signatures of positive selection along branch 55 (Fig. 19) identified under the various models (Table 12; Table 13; Fig. 19; Fig. 20) and were predicted to be surface exposed. LysM2, homology model of *Lj* Gifu-NFR5 LysM2 domain; KD, homology model of *Lj* Gifu-NFR5 kinase domain.

80
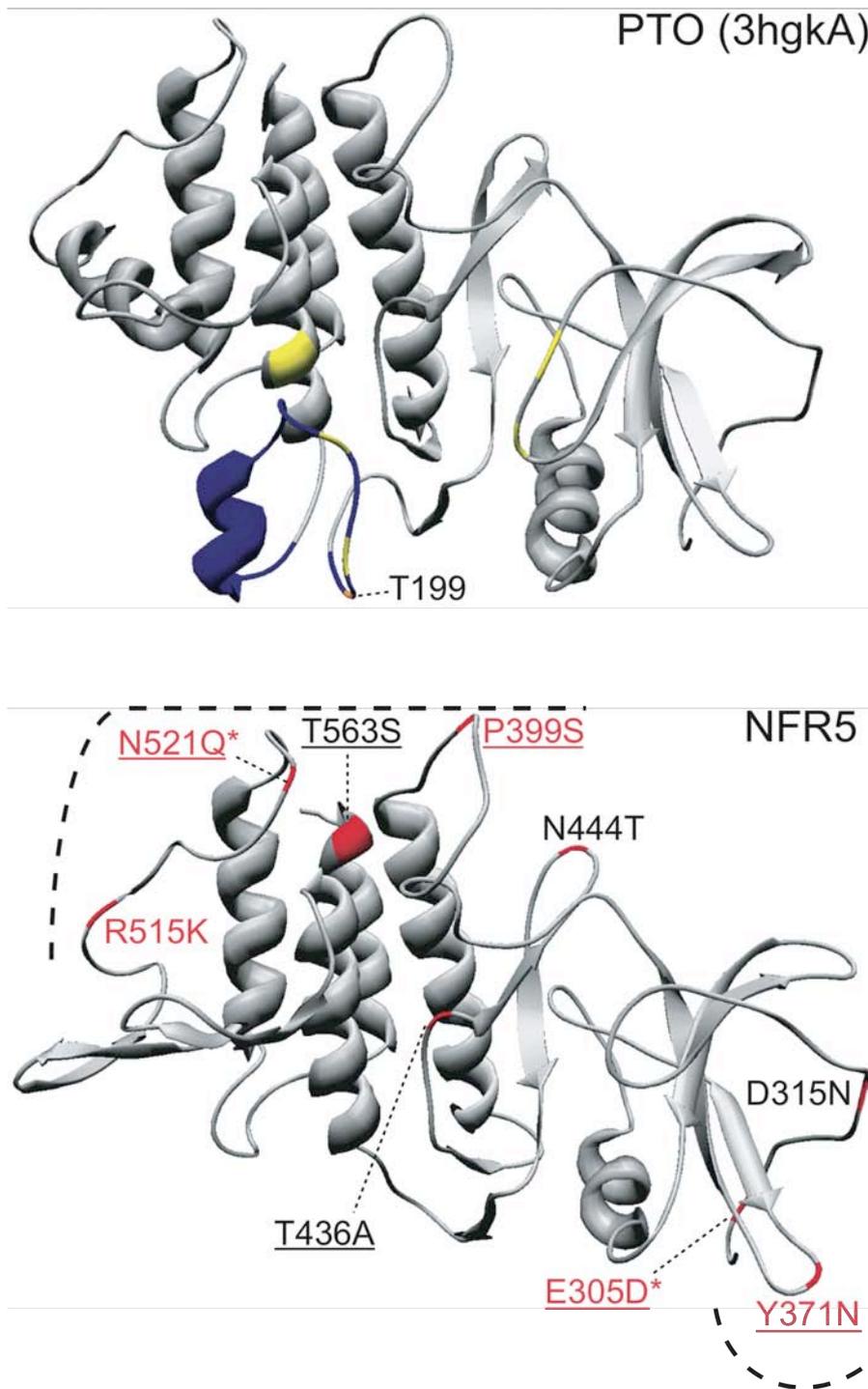
Figure 23. Pto kinase structure and homology model of *Lj* Gifu-NFR5 kinase domain.
The crystal structure of tomato Pto (*Pseudomonas syringae* pv. *tomato*) resistance (R) protein (PDB ID: pdb3hgk; Dong *et al.*, 2009) served as template for the homology model of *Lj* Gifu-NFR5 kinase domain. Blue, the activation segment (P+1 loop) is present in regular serine/threonine kinases and includes autophosphorylation site (T199), but is absent in the kinase domain of NFR5. Yellow, mutations at these site lead to a constitutive gain-of-function phenotype, which is an AvrPto independent immune response (Rathjen *et al.*, 1999; Wu *et al.*, 2004; Bernal *et al.*, 2005; Xing *et al.*, 2007). From a total of 13 sites, which distinguish the *NFR5* alleles *Lj* Gifu and *L. pedunculatus* (K0734_47F) in the kinase domain, nine could be placed onto the homology model. The remaining four sites were outside the region available for threading. Red ribbon, variable sites between *L. pedunculatus* (K0734_47F) and *L. japonicus* Gifu; red lettering, sites with signatures of positive selection; underlined lettering, sites varying uniquely between 'corniculatus' clade and 'pedunculatus' clade (Fig. 19); asterisks, sites with signatures of strong positive selection (high posterior probability values, Fig. 20).

IV.3 A quick T90 reporter line based assay to test the symbiotic activity of *Lotus japonicus* Gifu nodulation and AM mutants

IV.3.1 <u>Identification of double homozygous T90 lines with symbiotic mutant background</u>

Two double homozygous lines, T90 x *symrk-10* and *nfr1-1* x T90, were generated during this study. A first approach in selecting double homozygous F2 individuals that originated from reciprocal crosses between the symbiosis mutants, *nfr1-1* and *symrk-10*, and the reporter line T90 was the inoculation with *M. loti*. Three weeks post inoculation the non-nodulating individuals were subjected to PCR-based zygosity tests. PCR-amplified wild-type and T90 derived GUS construct alleles segregated in the non-nodulating F2 generation (Fig. 24a). Nine F2 individuals with a single amplicon at 1.3 kbp, corresponding to the T90 derived insertion of the GUS construct, were entered into the ZopRA Plant and Seed Database and tested for the presence of the symbiosis mutant alleles. The electropherograms of the PCR-amplified *NFR1* and *SYMRK* genes (Fig. 24b) show segregation of the mutant alleles *nfr1-1* and *symrk-10* in the selected T90-homozygous F2 generation. All identified single and double homozygous individuals were propagated to produce F3 progeny.

To determine if the symbiotic T90 GUS activity in the transgenic T90 lines with symbiotic mutant background is induced upstream or downstream of the respective mutant allele, the plants were treated with 10^{-8} M NF and two rhizobial stains (Fig. 25). After one week the GUS activity was assessed. The blue staining corresponds to symbiotic T90 GUS activity induced by the respective rhizobial strain or by NFs. In contrast to the T90 positive control no GUS staining was detected in the transgenic T90 lines with symbiotic mutant background. These results demonstrate that the symbiotic GUS activity is impaired in T90 lines with homozygous symbiosis mutant background.
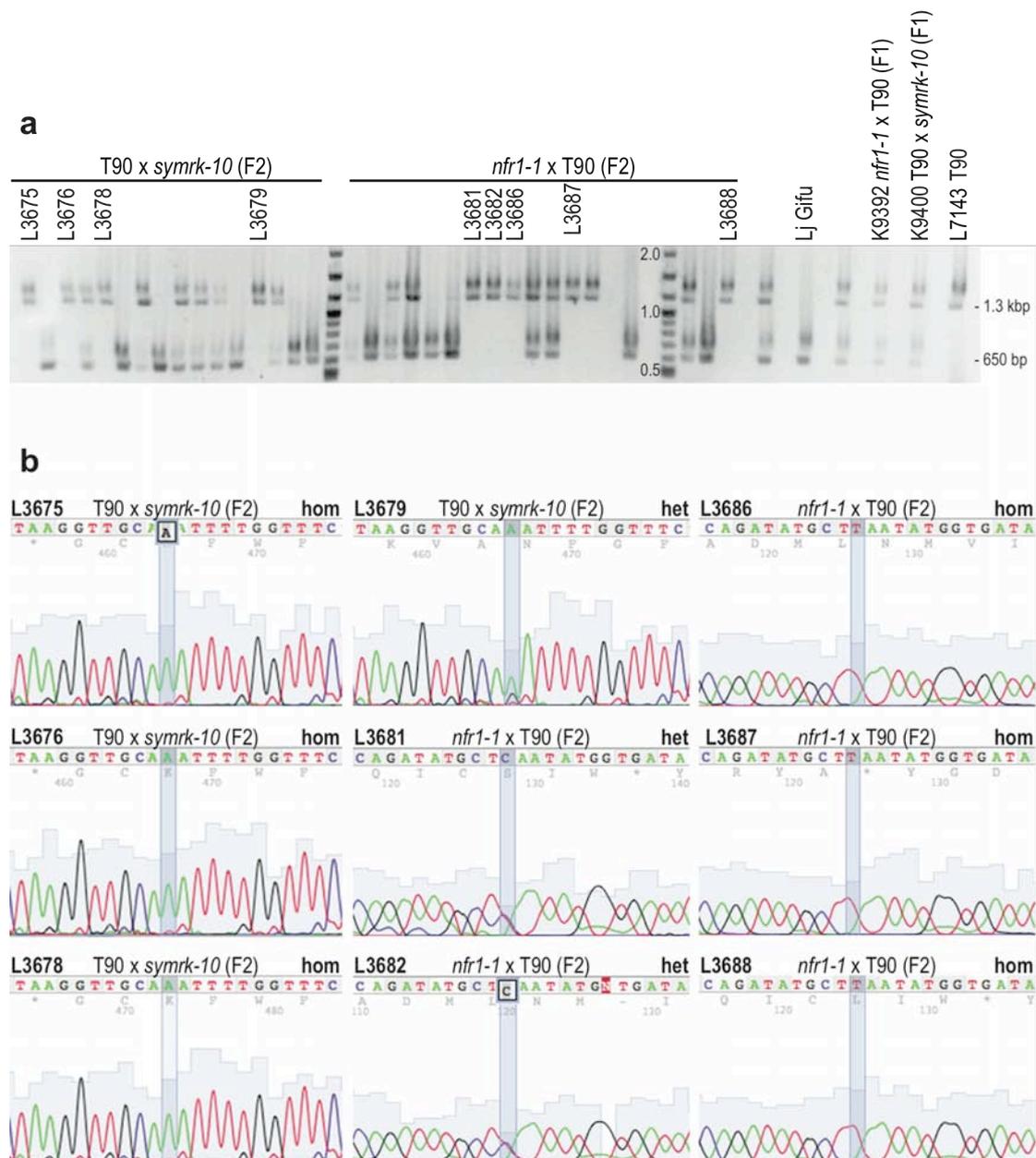
Figure 24. Zygosity tests of F2 individuals originating from crosses between symbiosis mutants and the *L. japonicus* Gifu GUS reporter line T90.
(a), PCR-based determination of the presence of the T90-derived insertion of the GUS construct. The wild-type *LjCbp1* gene gives rise to a product of 650 bp, the T90 *LjCbp1* gives rise to a 1.3 kbp product. (b), electropherograms of the PCR-amplified *symrk-10* and *nfr1-1* mutant gene regions, including the site of the mutation (shaded). Plant line designations correspond to the ZopRA Plant and Seed Database nomenclature; numbering (e.g. L3675) corresponds to mother plant. hom, homozygous; het, heterozygous.

83
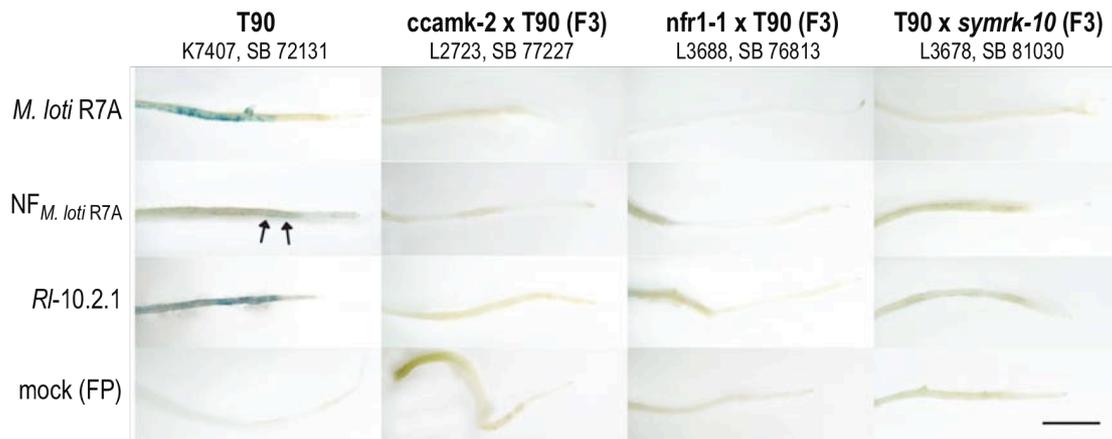
Figure 25. GUS induction test of double homozygous T90 lines with symbiotic mutant background and reporter line T90 as positive control.
The β-glucuronidase activity (GUS) was assessed by histochemical staining with X-Gluc one week post application of NF (10^{-8} M) or inoculation with rhizobia or mock inoculation. Plant line designations correspond to the ZopRA Plant and Seed Database nomenclature; numbering (e.g. L3675) corresponds to mother plant. *Rl*, *Rhizobium* cf. *leguminosarum*; arrow, weak GUS staining. Scale bar = 500 μm.

IV.3.2 Symbiotic GUS activity of single and double homozygous T90 lines with symbiotic mutant background induced by selected rhizobium strains

Different rhizobium strains were found to take various infection routes. We tested three rhizobium strains, which differ in their host ranges, for their potential to induce symbiotic GUS activity in single and double homozygous lines *nfr1-1* x T90 and T90 x *symrk-10* (Fig. 26a-d). At an early time point (1 wpi) the strains *M. loti* and *Bradyrhizobium* sp. induced strong GUS activity in the reporter line T90 and in progeny of the single homozygous T90 lines with symbiotic mutant background (Fig. 26a,c). The GUS activity segregated in a nearly 3:1 ratio among progeny of single heterozygous individuals of lines T90 x *symrk-10* and *nfr1-1* x T90 (Fig. 26a,b).

The strain *Rhizobium* cf. *leguminosarum* 10.2.1 (hereafter *Rl*-10.2.1) induced weak or strong GUS activity in T90 individuals. Interestingly, this strain also was able to induce weak GUS activity in some of the double homozygous individuals. This weak GUS staining was more profound after an elongated staining period of 48h instead of 24h. Also, some earlier GUS⁻ rated individuals were then reappraised GUS⁺. On some double homozygous plants *M. loti* and *Bradyrhizobium* sp. induced a blue staining that was spatially restricted to one or few rhizodermal cells (Fig. 26a, stars; Fig. 26c, arrows).

Overall, at a later time point (3 wpi) both, the *M. loti* and the *Bradyrhizobium* sp. induced GUS activity pattern were strong in areas of primordia or nodules and had vanished in the rest of the root (Fig. 26b,d). Surprisingly, one bump was found on one

plant of the *M. loti* inoculated double homozygous line T90 x *symrk-10* (L3678). This blue stained bump indicates symbiotic GUS activity (Fig. 26d, arrow). Despite the similar GUS induction pattern, the nodulation phenotypes induced by *M. loti* and *Bradyrhizobium* sp. vary. On the T90 control *M. loti* induced a lower number of nodules than *Bradyrhizobium* sp., but these nodules were bigger and pink (Fig. 26d,e). All plants, which were inoculated with *Rl*-10.2.1 were GUS⁻ at 3 wpi, however, on the single homozygous line *nfr1-1* x T90 (L3682) a bump formed (Fig. 26d, star).
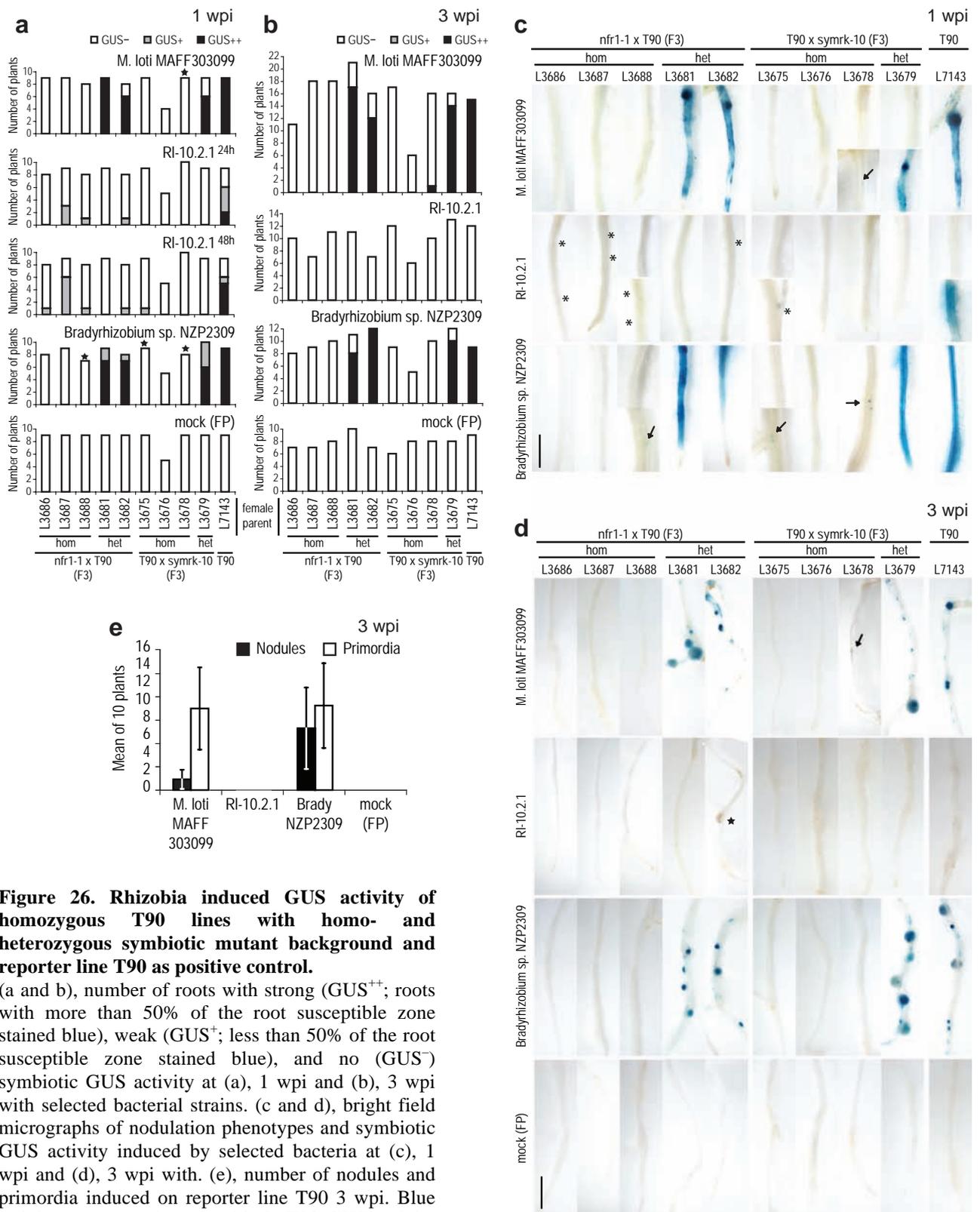
Figure 26. Rhizobia induced GUS activity of homozygous T90 lines with homo- and heterozygous symbiotic mutant background and reporter line T90 as positive control.

(a and b), number of roots with strong (GUS^{++}; roots with more than 50% of the root susceptible zone stained blue), weak (GUS^{+}; less than 50% of the root susceptible zone stained blue), and no (GUS^{-}) symbiotic GUS activity at (a), 1 wpi and (b), 3 wpi with selected bacterial strains. (c and d), bright field micrographs of nodulation phenotypes and symbiotic GUS activity induced by selected bacteria at (c), 1 wpi and (d), 3 wpi with. (e), number of nodules and primordia induced on reporter line T90 3 wpi. Blue staining corresponds to symbiotic T90 GUS activity induced by the respective bacterial strain. Plant line designations correspond to the ZopRA Plant and Seed Database nomenclature; numbering (e.g. L3675) corresponds to mother plant. Asterisk, area of weak blue staining on GUS^{+}-roots; arrow, one to few blue-stained cells on GUS^{-}-roots; star, bump; *Rl*, *Rhizobium* cf. *leguminosarum*. Scale bars = 1 cm.

86

IV.3.3 <u>Double homozygous T90 lines with symbiotic mutant background expressing the T265D gain-of-function construct of CCAMK</u>

T265D, the gain-of-function construct of CCAMK was over-expressed on hairy roots from four double homozygous T90 lines with symbiotic mutant background. This construct induced GUS activity in hairy roots of all T90 lines with symbiotic mutant background as well as on the reporter line T90 (Fig. 27). Non-inoculated hairy roots transformed with the empty vector control were GUS⁻ (Fig.27).

The GUS activity in hairy roots also was tested one week post inoculation with rhizobium strains *M. loti*, *Rl*-10.2.1, and *Bradyrhizobium* sp. Independent of the mutant allele and the rhizobium strains GUS activity was induced in hairy roots transformed with the T265D construct and in hairy roots of reporter line T90 transformed with the empty vector control (Fig. 27). No GUS activity was detected in the T265D transformed hairy roots of line *ccamk-2* x T90 inoculated with *Bradyrhizobium* sp. (Fig. 27). This GUS⁻ phenotype was probably caused by the strongly reduced germination frequency and growth performance of this line.

The mock inoculated hairy roots of T90 x *symrk-10* transformed with the T265D construct displayed no RHC, while the application of *M. loti* induced RHC and the formation of infection pockets (Fig. 28, insets). At four weeks post inoculation the T265D expressing hairy roots of the reporter line T90 and T90 x *symrk-10* line formed ITs (Fig. 29a, insets). These results demonstrate that the infection deficiency of *symrk-10* was abolished in the hairy roots transformed with T265D gain-of-function construct of CCAMK.

At four weeks post mock inoculation the T265D expressing hairy roots of the reporter line T90 and T90 x *symrk-10* line formed bumps and small nodules (Fig. 29b). Bigger nodules formed on hairy roots inoculated with *M. loti*. The nodules also were red fluorescent, confirming that they were infected with *M. loti* expressing DsRed.
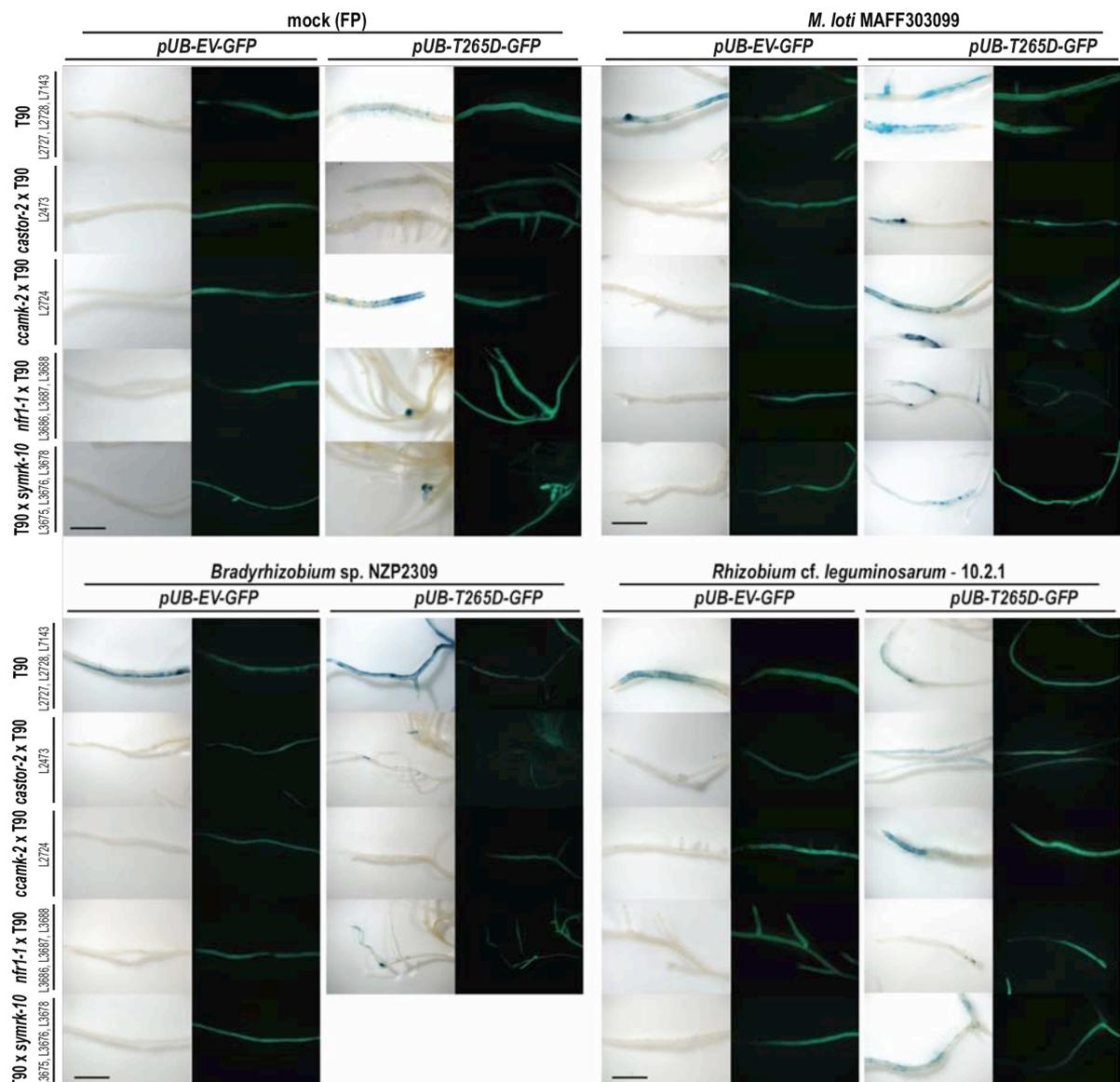
Figure 27. GUS induction in hairy roots of double homozygous T90 lines with symbiotic mutant background and reporter line T90 as positive control at 1 wpi with selected rhizobial strains.
The plants were transformed with AR1193 carrying an empty vector control (pUB-GW-EV-GFP) or T265D, an autoactive version of *ccamk* (pUB-GW-T256D-GFP) and mock inoculated r inoculated with *M. loti*, *Rhizobium* cf. *leguminosarum*, or *Bradyrhizobium* sp. The panels show bright field (1st and 3rd column) and green fluorescent protein (GFP, 2nd and 3rd column). Plant line designations correspond to the ZopRA Plant and Seed Database nomenclature; numbering (e.g. L3675) corresponds to mother plant. Asterisk, area of weak blue staining. Scale bar = 1 cm.
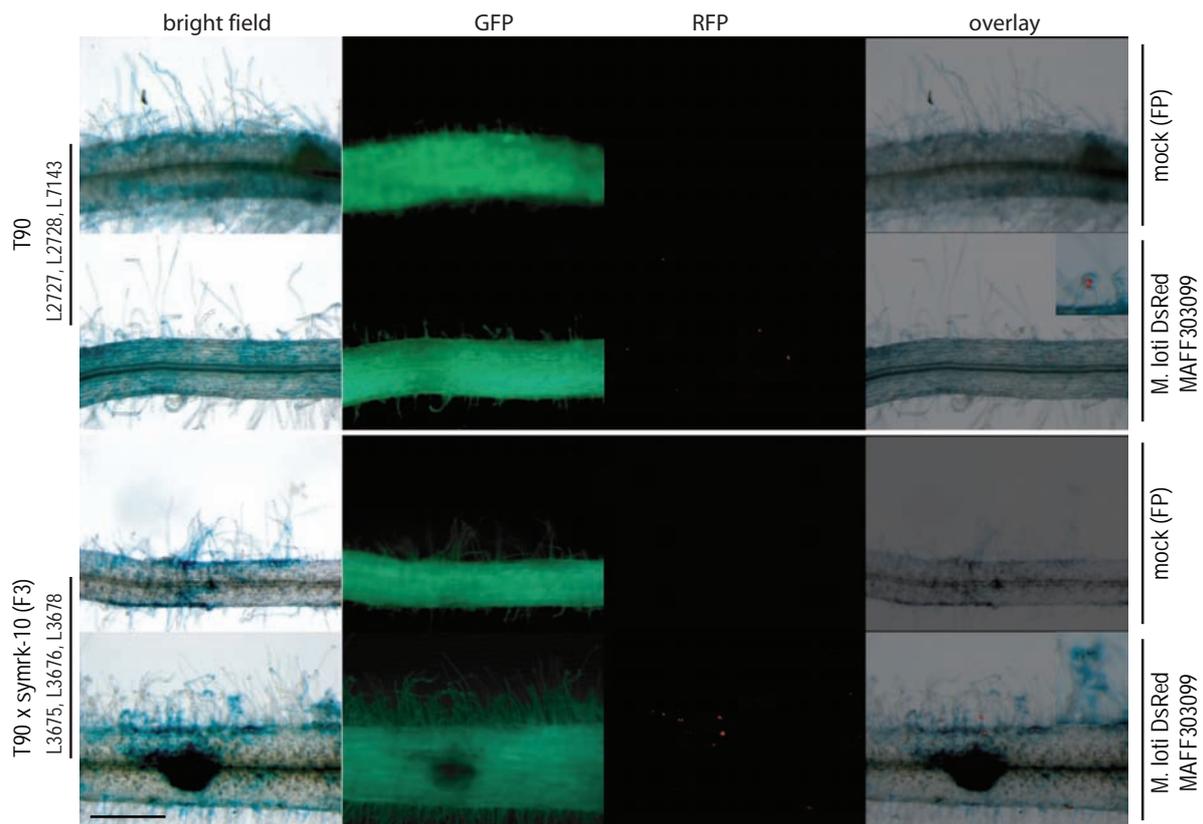
Figure 28. Infection phenotypes and GUS activity of hairy roots from double homozygous T90 *x symrk-10* line and reporter line T90 as positive control at 1 wpi with *M. loti*.

Hairy roots expressing pUB-GW-T265D-GFP or the empty vector control pUB-GW-EV-GFP. The panels show bright field (bf), green fluorescence protein (GFP), red fluorescence protein (RFP, DsRed) labeled rhizobia, and the combined image of bright field and RFP signal (overlay). Inset, magnifications of framed area. Plant line designations correspond to the ZopRA Plant and Seed Database nomenclature; numbering (e.g. L3675) corresponds to mother plant. Scale bar = 500 µm.
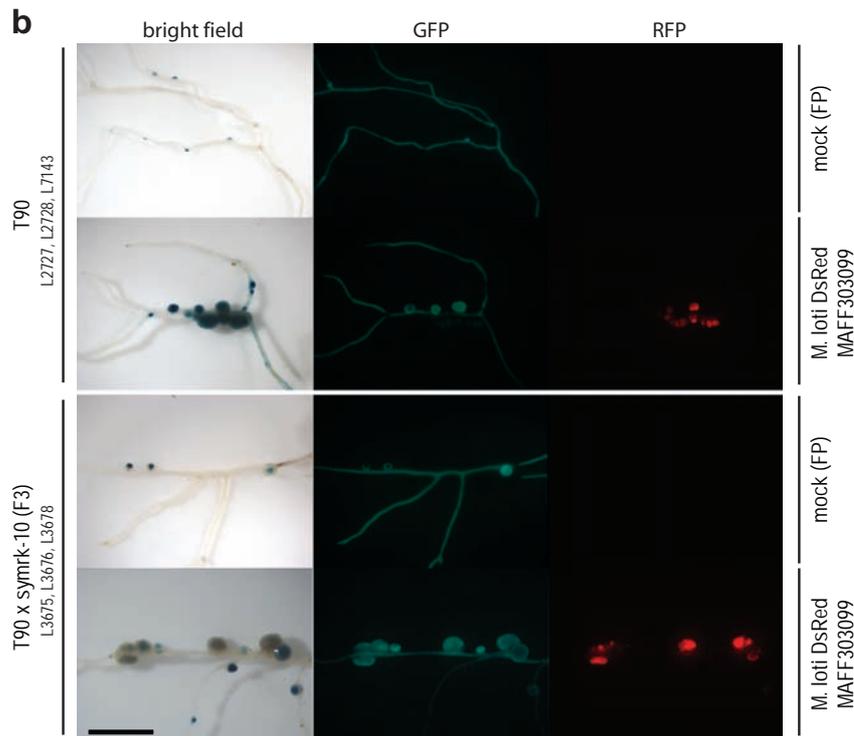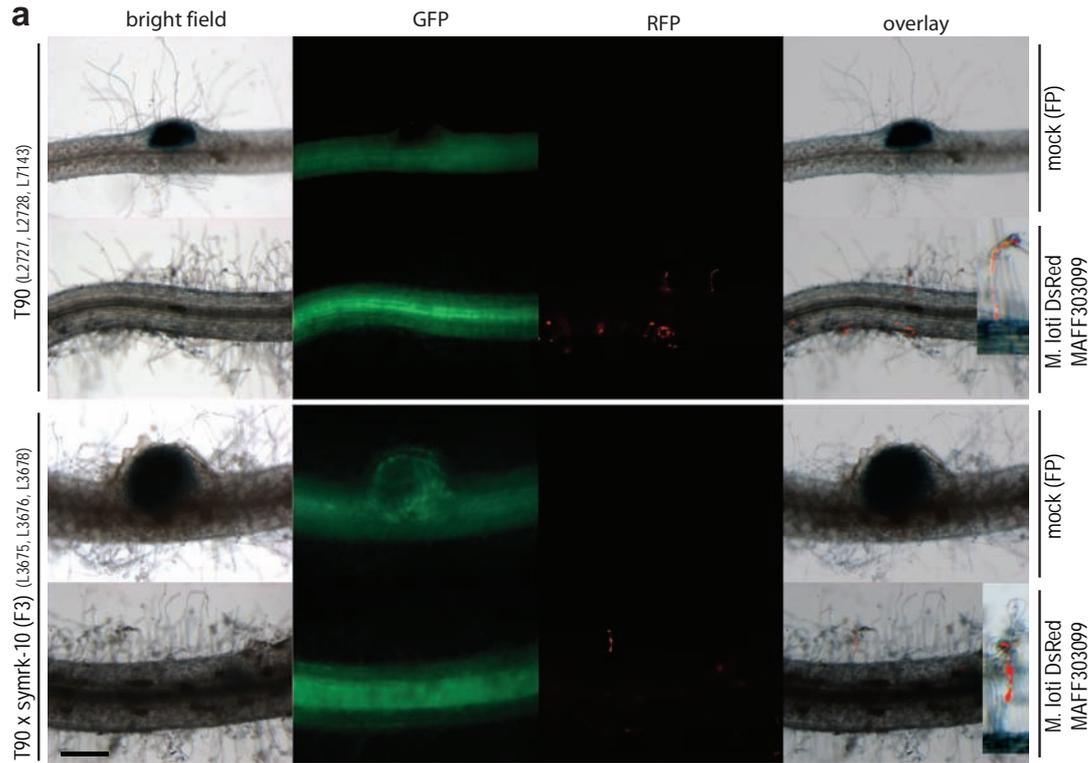
Figure 29. Symbiotic phenotypes and GUS activity of hairy roots from double homozygous T90 *x symrk-10* line and reporter line T90 as positive control at 4 wpi with *M. loti*.
(a) Infection phenotypes and (b), nodulation phenotypes of hairy roots expressing pUB-GW-T265D-GFP or hairy roots expressing the empty vector control pUB-GW-EV-GFP. The panels show bright field (bf), green fluorescence protein (GFP), red fluorescence protein (RFP, DsRed) labeled rhizobia, and the combined image of bright field and RFP signal (overlay). Inset, magnifications of framed area. Plant line designations correspond to the ZopRA Plant and Seed Database nomenclature; numbering (e.g. L3675) corresponds to mother plant. Scale bar = 500 µm (a); 1 cm (b).

IV.3.4 Gene induction in vacuum infiltrated and co-cultivated plants

We tested if the T265D gain-of-function construct of CCAMK induces T90 GUS activity upon vacuum infiltration of young seedlings (Fig. 30a,b). Different conditions were tested to eliminate the possibility of aberrant GUS activity. We used two negative controls to test for a possible AR1193 related induction of GUS. One control was the mock infiltration of plants with VIM and subsequent mock inoculation with FP medium, the other control was the infiltration of the plants with AR1193 carrying the empty vector and subsequent mock inoculation with FP medium. Most of the mock infiltrated plants showed no GUS activity (Fig. 30a). Only one plant per line, except for *castor-2* x T90 plants, displayed GUS activity. However, this faint blue staining was spatially restricted to small areas in the root (Fig. 30b, arrows).

Surprisingly, the empty vector infiltrated mock inoculation controls of T90 were all GUS$^+$, while of the T90 x *symrk-10*, *nfr1-1* x T90, and *castor-2* x T90 plants some were GUS$^+$ and some were GUS$^-$ (Fig. 30a). Under these conditions a wide area of the GUS$^+$ roots was moderately stained blue (Fig. 30b). Interestingly, all *ccamk-2* x T90 plants were GUS$^-$. Almost all empty vector infiltrated roots that were inoculated with *M. loti* were moderately stained blue, and only two T90 x *symrk-10* plants were GUS$^-$ (Fig. 30a,b).

The infiltration with the T265D gain-of-function construct of CCAMK led to GUS activity in all or in the majority of T90 individuals with symbiotic mutant background except for T90 x *symrk-10*. Here eight mock inoculated seedlings were GUS$^-$ and two were GUS$^+$.

Generally, the quality of the GUS staining varied from a faint staining in the mock infiltrated plants to a moderate staining in the AR1193 infiltrated plants independent of the construct that was used. The infiltrations with *A. rhizogenes* strain AR1193 enhanced the GUS activity even in the different mutant backgrounds.

91
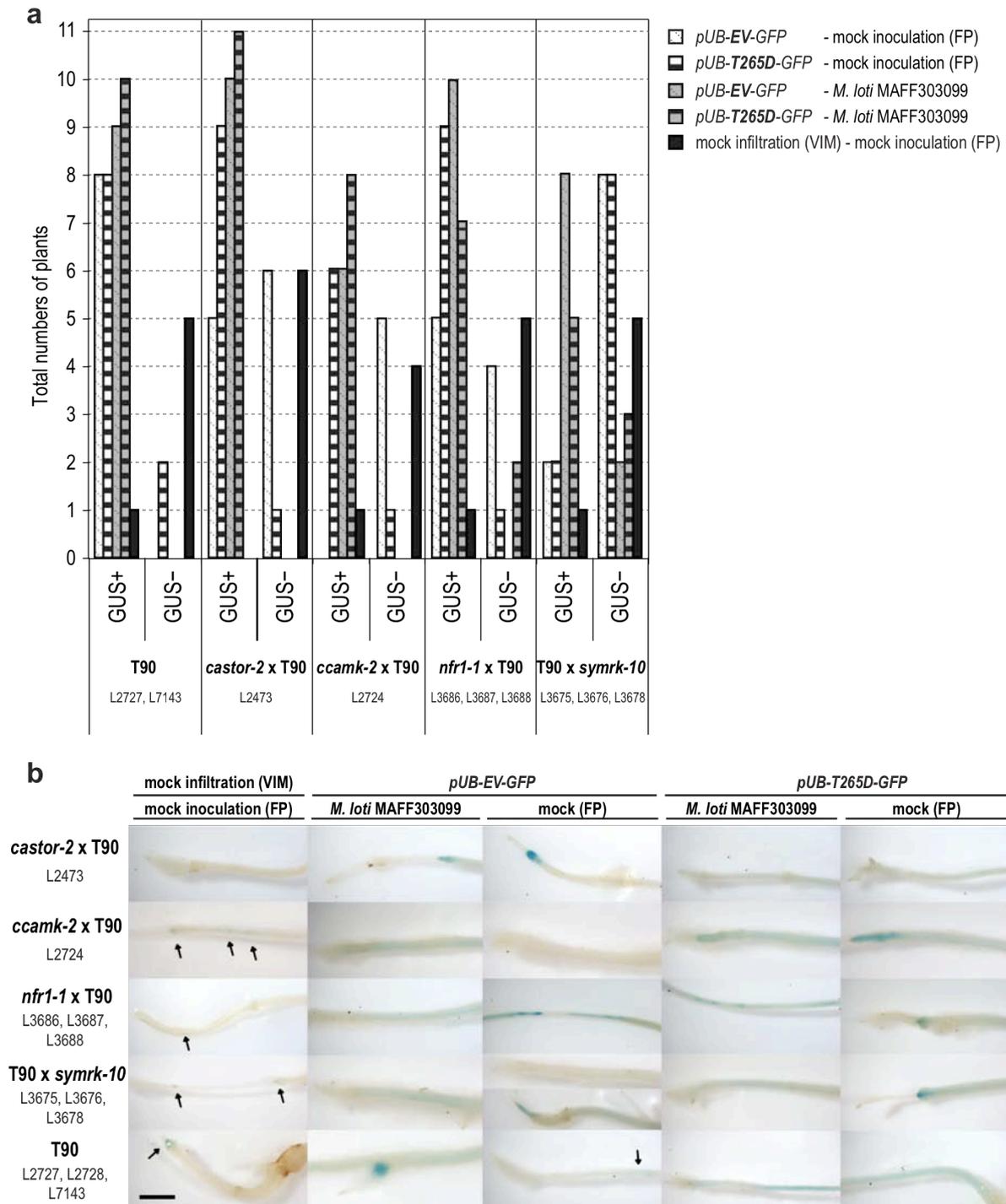
Figure 30. GUS activity of F3 seedlings after transient transformation with AR1193 and 4 dpi with *M. loti*.
(a), number of roots stained blue (GUS⁺); number of roots without blue staining (GUS⁻); (b), bright field micrographs of GUS stained roots. Seedlings of double homozygous T90 lines with symbiotic mutant background and reporter line T90 as positive control were vacuum infiltrated with AR1193 carrying either the T265D gain-of-function construct of CCAMK (pUB-GW-T265D-GFP) or the empty vector control (pUB-GW-EV-GFP). As negative control seedling were mock infiltrated with vacuum infiltration medium (VIM) only. The GUS activity of the infiltrated seedlings was assessed 4 dpi with *M. loti* or mock inoculation. Plant line designations correspond to the ZopRA Plant and Seed Database nomenclature; numbering (e.g. L3675) corresponds to mother plant. Scale bar = 500 µm (b).

V DISCUSSION

V.1 Distribution of genetic variation and host specificity among nitrogen fixing symbionts from closely related *Lotus* species

V.1.1 <u>Collection data reveal a lower stringency for restricting incompatible bacteria in *L. corniculatus*</u>

The fundamental work describing compatibility groups in *Lotus* that was earlier performed previously is reviewed in Saeki & Kouchi (2000). Our data largely confirms the presence of a very robust compatibility barrier between *L. pedunculatus* and *L. corniculatus*. We found that under natural conditions the bradyrhizobial strains are associated with *L. pedunculatus*, while *Mesorhizobium* forms efficient nodules with *L. corniculatus*. The two newly isolated bradyrhizobial strains that are closely related strains of *Rhodopseudomonas* sp. both form fully efficient symbioses with *L. pedunculatus*. However, their nodulation capacity on *L. japonicus* varies greatly, between Nod⁻ (strain 21.3B.2) and Nod⁺Fix⁻ (strain 21.4C.2).

However, *Lotus corniculatus* had less ability to restrict nodule occupancy by suboptimal or incompatible rhizobial strains under both natural and experimental conditions. For example, from a single *L. corniculatus* nodule we co-isolated non-compatible *Rl*-10.2.1 and compatible *Mesorhizobium* (10.2.2). The latter induced fully efficient nodules on *L. japonicus, L. burttii, L. filicaulis, L. corniculatus* and *L. glaber*.

V.1.2 <u>New infection and nodule organogenesis polymorphisms between *Lotus* accessions revealed through rhizobial isolates</u>

Strain *Rl*-10.2.1 induced a range of incompatible nodulation phenotypes, from Nod⁻ on *L. filicaulis* and intercellular bacterial entry correlated with EIZ on *Lj* Nepal to the formation of infected, yet inefficient nodules on *L. burttii*.

The symbiotic capacities of the tested *Lotus* species varied greatly upon inoculation with different rhizobia. Generally, *L. filicaulis* is strongly restricted to fully compatible symbionts in the genus *Mesorhizobium* and late incompatible symbiont *Bradyrhizobium*. In contrast, *L. burttii* nodulation is not limited to fully compatible rhizobia. For example, *Rl*-10.2.1, *Rlv*-DZ, *Rlv*-DZL, and NGR234 are Nod⁻ on *L. filicaulis*, but they induce infected nodules on *L. burttii*. However, these

nodules are inefficient.

Cross-inoculation experiments revealed a high level of variation among the tested *Lotus* species in their symbiotic compatibility at stages post organogenesis. Bradyrhizobial strains and induce white, inefficient nodules on *L. japonicus, L. burttii, L. filicaulis, L. corniculatus,* and *L. glaber*. Similarly, strain NGR234 shows an initial fixation deficiency on *L. japonicus* and *L. burttii*.

V.1.3 <u>Infection modes vary with *Lotus* genotype</u>

Legumes have evolved different nodule morphologies and rhizobial entry mechanisms summarized in Sprent (2007). In a recent study with spontaneous nodulating double and triple mutants of *L. japonicus*, nodule organogenesis was successfully uncoupled from infection and all three alternative entry processes described in legumes were possible on *L. japonicus* (Sprent, 2007; Madsen *et al.*, 2010). These three alternative host controlled entry processes are: 1) intercellular infection, 2) crack entry with or without ITs, and 3) root hair infection. The order of these three modes of infection may correspond to the timing of their evolutionary origin with bacterial intercellular infection without ITs as the earliest form of infection and progressing to the more derived state involving root hair infection and intracellular accommodation.

Some host species express two infection modes, depending on the environmental conditions. For example, nodules can form on both stems and roots of the semi-aquatic *S. rostrata*, and this species can switch between infection pathways from epidermal root hair invasion under non-flooded conditions to cortical intercellular invasion at lateral root bases via so-called "cracks" in the epidermis under flooded conditions (Goormachtig *et al.*, 2004). Likewise, the wetland plant *L. uliginosus* is capable of root hair nodulation as well as infection via enlarged epidermal cells and cortical IT-formation (James & Sprent, 1999). We discovered that *Lj* Nepal can switch from slow intercellular bacterial entry with *Rl*-10.2.1 to fast track root hair IT infection with compatible *M. loti*. This result demonstrates that a single *Lotus* genotype can be infected via different modes depending on bacterial species.

V.1.4 <u>Stringency of NF recognition varies for different infection routes</u>

In *Sesbania*, the NF structural requirements are more stringent for the root hair infection than they are for the more ancient mode of intercellular invasion. This is in accordance with our observations that *Rlv*-DZ can infect *Lj* MG20 via enlarged epidermal cells, but is blocked from further passage into the root cortex, while normal infection is recovered with *Rlv*-DZL.

Based on earlier studies of *Medicago* sp. with bacterial mutants of *S. meliloti* 2011 a model was proposed, with a 'signaling receptor' for induction of early responses with low structural requirements of NFs and an 'entry receptor' that controls infection and requires more stringent matching of NF structure (Ardourel *et al*., 1994). In *L. japonicus* two receptors are known to control bacterial infection, NFR1 and NFR5 (Madsen *et al*., 2003; Radutoiu *et al*., 2003). In *M. truncatula* the NFP protein, the *Lj*-NFR5 ortholog, acts like a NF signaling receptor and an additional LysM domain-containing receptor-like kinase (LYK3) specifically controls the bacterial entry via ITs in an NF dependent manner (Limpens *et al*., 2003; Smit *et al*., 2007). Generally, our results are in accordance with the proposed two-receptor model, with different levels of NF structure stringency demands. The different entry pathways may require different receptors and/or different NF structural stringencies. Legumes have a large set of LysM domain containing RLK (Lohmann *et al*., 2010), and some are specifically expressed upon RNS. Members of this gene family could be involved in switches of the infection pathways in a NF-structure dependent manner and/or function sequentially from the epidermis towards the root cortex.

V.1.5 <u>Evidence for negative regulatory mechanisms controlling nodulation and infection in *Lotus*</u>

The *Rlv*-D induced Nod⁻ phenotypes on *L. filicaulis* could be partially dependent on NF-signaling via the NFRs (Radutoiu *et al*., 2007). Assuming that the different infection pathways are controlled by different LysM-RLKs it is surprising that we found no signs of non-root hair infection on *L. filicaulis*. Either the alternative intercellular pathway does not exist in *L. filicaulis* or it has stringent NF-structure requirements. An alternative explanation could be that *L. filicaulis* efficiently excludes incompatible symbionts via an additional dominant pathway repressing

organogenesis and infection, which upon activation, could lead to incompatibilities and abortion of the RNS.

Such negative regulatory pathways would also explain the strict exclusions of incompatible rhizobia *M. loti* NZP2213 by *L. pedunculatus*, where organogenesis is induced but only non-infected nodules form (Pankhurst *et al.*, 1979). There is even evidence for a defense response via flavolans produced by *L. pedunculatus*. The strain NZP2213 is sensitive to the plant produced prodelphinidin, which probably leads to the Nod$^+$Fix$^-$ phenotype (Pankhurst & Jones, 1979). It has also been shown that *L. pedunculatus* increases the production of certain flavonoids only in non-fixing nodules (Cooper & Rao, 1992).

Moreover we found that different patterns of GUS expression of the T90 reporter line correlated with different NF-structures produced by variants of RBL5560. *Rlv*-D, which produces unfucosylated NFs, induces weak GUS activity. Two derivatives of RBL5560 that produce NFs with either an acetylated or non-acetylated fucosyl residue on C-6 of the reducing-terminal GlcNAc both induce fast and strong GUS activity and organogenesis on T90 plants. This pattern of fast and strong GUS activity typically coincided with nodule organogenesis. Both *Bj* USDA110 Δ132 and the wild-type fail to induce nodules on *Lotus japonicus* (Göttfert, personal communication, September 21, 2010). Curiously, Δ132 induced fast and strong GUS activity. This observation suggests a matching NF-receptor pair in this combination, and calls for the presence of a negative regulatory pathway that repress nodulation by this strain.

V.1.6 Compatibility during late nodule organogenesis and infection

In our experiments, the growth benefits to the plants resulting from the symbiosis of a single rhizobial strain varied between different hosts. For example, *Brady*NZP2309 induced infected nodules on all *Lotus* species tested, but effective nodules formed only on *L. pedunculatus*. While the infection process has been studied extensively, little is known about molecular compatibility of symbiotic partners during later stages of the nodulation process (Den Herder & Parniske, 2009).

Rhizobium sp. NGR234 induces *M. loti*-like infections and organogenesis on three ecotypes of *L. japonicus* (MG20, Gifu, and Nepal) and on *L. burttii*, but with a delay in N-fixation efficiency (Schumpp *et al.*, 2009). The delay in N-fixation efficiency can be abolished on *L. burtii* and partially on *Lj* Nepal by an EPS mutant of

NGR234 (NGRΩ*exoK*). This mutant produces EPS devoid of low-molecular-weight forms. Through co-inoculation experiments with EPS-deficient *R. leguminosarum* RBL5833 and inoculations with several transconjugants, EPS was shown to contribute to host-plant specificity of *Vicia sativa* nodulation. But it is not the main specificity determinant. The wild-type strain RBL5833 induced abortive ITs on *V. sativa* subsp. *nigra* (Laus & van Brussel, 2005). These studies, taken together with our results on various *Lotus* species, suggest the action of host specific perception of EOS during later stages of the RNS. The absence of EOS in the mutant NGRΩ*exoK* positively affects the symbiosis with *L. burttii* but has no effect in the RNS with *L. japonicus* Gifu six weeks after inoculation. Further detailed phenotypic analysis are needed to identify effects of the EOS triggered N-fixation deficiency. The species and ecotype specific differences in N-fixation phenotypes with NGRΩ*exoK* could help to identify loci conferring specificity during late stages of root nodule development and infection.

Taken together, we identified and described early and late compatibility phenotypes in natural populations of European *Lotus* species. These *Lotus* species with contrasting phenotypes are crossable to each other, and for some of them recombinant inbred lines are available, which will facilitate identification of new loci encoding or contributing to these naturally evolved compatibilities of both symbiosis partners.

V.2 **Molecular evolution of Nod Factor Receptor 5 correlates with contrasting rhizobium specificity of European populations of *L. pedunculatus* and *L corniculatus***

V.2.1 <u>Evolutionary history of the *NFR5* gene</u>

We investigated the infection specificity and the nodulation compatibility of the parental lines, *L. pedunculatus* and *L. japonicus* Gifu, and in-vitro clones of the F1 hybrids hereof, upon cross-inoculations with *Bradyrhizobium* sp. and *M. loti.* We show that incompatibility can arise at multiple developmental steps during the initiation of RNS on *Lotus* host species. The infection incompatibility of *M. loti* and *L. uliginosus* together with the induction of developmental processes including primordium organogenesis suggests a certain level of rhizobial NF recognition by the plant. Based on the proposed two-receptor model we can assume that the entry of *M.*

loti on *L. pedunculatus* requires a higher structural stringency of the NF (Ardourel *et al.*, 1994). Observations in *S. rostrata*, where root hair infection required a more stringent NF structure than crack entry support the hypothesis of additional epidermal or cortical NFRs (D'Haeze *et al.*, 2000; Goormachtig *et al.*, 2004). In the case of F1 hybrids one or more alleles inherited from *L. japonicus* Gifu confer the hybrids infection compatibility at later cortical infection stages.

Furthermore, we used a candidate gene approach to study the evolutionary history of one of the genetic determinants of species discrimination in *Lotus* species, the *NFR5* gene (Madsen *et al.*, 2003; Radutoiu *et al.*, 2007). We discovered high levels of amino acid divergence between host species differing in symbiotic compatibilities, concentrated in the LysM2 domain and the region between the LysM1 and LysM2 domain, which are implicated in NF binding and recognition specificity. Interestingly, we also found higher levels of amino acid divergence in the central and the C-terminal region of the kinase domain. Furthermore, using ancestral state reconstruction, we uncovered a number of amino acid substitutions that are associated with the differences in symbiont compatibility. 40.9 % of these substitutions occurred in the LysM domains, including the previously described K118L substitution, which is associated with a switch in nodulation specificity of *L. filicaulis* towards *Rlv*-DZL(Radutoiu *et al.*, 2007). The rate of non-synonymous evolution in the LysM domain is 2.73 times that of the rest of the gene and although only 5% of the gene shows an omega value >1 (indicative of positive selection), 21.4 % of these targets of adaptation occur in the LysM domains. Taken together, our population and evolutionary genetic analyses indicate that purifying selection is the predominant evolutionary force acting at the *NFR5* gene. However, a limited number of positions do show elevated amino acid divergence between species differing in recognition specificity and nearly all of these positions are predicted to be exposed residues of the LysM domains or the kinase domain.

V.2.2 <u>Plant controlled selection and exclusion are active in RNS</u>

From an evolutionary genetic perspective, *NFR5* appears to be a target of adaptive evolution involved in the discrimination between potential nodulating bacterial symbionts. However, our phenotypic studies, combined with the use of a battery of informative mutant bacterial strains, implicate additional genetic components encoded by both the host and symbiont involved in host-symbiont compatibility. In

particular, *L. pedunculatus* shows a strong incompatibility to infection by *Meso*MAFF303099, which forms efficient RNS with *L. japonicus*. F1 hybrids from a cross between *L. pedunculatus* x *L. japonicus* are initially compatible with *Meso*MAFF303099, but become incompatible later during infection at the stage in which the bacteria would normally be released into the host cells from the IT. The early infection compatibility, in which ITs are formed, is transmitted as a dominant trait from the *L. japonicus* parent, while the termination of infection at the stage of bacterial uptake, is transmitted as a dominant trait from the *L. pedunculatus* parent. The termination of the symbiosis at this later stage can be regarded as a second checkpoint in the host-symbiont interaction. The strain *Meso*DT3S can cross this late infection barrier and forms wild-type nodules on the hybrid *LpLj*6.2 suggesting that incompatibility is dependent upon an intact T3SS and probably one or more type III secreted effector proteins.

*Meso*MAFF303099 is typically unable to infect *L. pedunculatus* and compatibility is initially possible in the hybrids via genetic contributions from the *L. japonicus* host, but breaks down at a later stage due to factors from *L. pedunculatus*, which are dependent upon a functioning T3SS. This implicates the T3SS as a determinant of incompatibility between *Mesorhizobium* and *L. pedunculatus*. However, since *Meso*DT3S cannot overcome the infection barrier in *L. pedunculatus* alone*,* but is able to enter the nodule upon co-inoculation with *Brady*NZP2309, further bacterial signals are implicated. It is likely that one or more signals, possibly NFs, are involved in the early infection incompatibility, and another signal, probably translocated via the T3SS, mediates the late infection blockage. Following the gene-for-gene hypothesis, the two infection barriers present in the host should be encoded by at least two loci. Our phenotypic studies of the F1 hybrids in which early IT progression is inherited as a dominant trait from *L. japonicus* parent, while the blockage of bacterial release from the IT is inherited as a dominant trait from *L. pedunculatus* parent, is consistent with at least two genetic factors inherited independently.

In contrast to the observation from cross-inoculations of *Mesorhizobium* on its non-host, *L. pedunculatus*, *Brady*NZP2309 can form nodules on *L. japonicus,* but the nodules are inefficient. *Brady*NZP2309 also forms functional nodules on the hybrid between *L. japonicus* and *L. pedunculatus*, in contrast to the observations from infections with *Mesorhizobium*, in which a T3SS dependent incompatibility encoded

by the L. *pedunculatus* parent prevented release of the bacteria into the hybrid cells. This indicates that putative allelic contributions encoding incompatibility from the *L. japonicus* parent act recessively in the hybrid. It is conceivable that *Brady*NZP2309 is lacking one ore more signals triggering further plant responses in the nodule, for example leghemoglobin production. Some rhizobial type III secreted effectors were shown to be specificity determinants, with either positive or negative effects on different hosts (Kambara et al., 2009; Sánchez et al., 2009). The pathogen *Ralstonia solanacearum,* carrying the symbiotic plasmid of *Cupriavidus taiwanensis,* was successfully converted into a *Mimosa*-nodulating and infecting symbiont by inactivation of structural genes of its T3SS (Marchetti *et al.*, 2010). Likewise, *Meso*DT3S is equally efficient on *L. japonicus* as the wild-type *Meso*MAFF303099*,* but in contrast to the wild-type, *Meso*DT3S can form efficient nodules on *L. halophilus* (Okazaki *et al.*, 2010), indicating that factors secreted via the T3SS inhibit nodule formation of *Mesorhizobium* on *L. halophilus*. In RNS effector proteins are recognized by Toll-interleukin receptor/nucleotide-binding site/leucine-rich repeat (TIR-NBS-LRR) class of plant resistance (R) proteins (Yang *et al.*, 2010). Two TIR-NBS-LRR class R protein genes from soybean were recently cloned, Rj2 and Rfg1 and demonstrated to restrict the nodulation with specific strains of *B. japonicum* and *Sinorhizobium fredii* (Yang *et al.*, 2010). This suggests a common recognition mechanism underlying symbiotic and pathogenic host-bacteria interactions.

L. pedunculatus also strictly excludes other *M. loti* strains. For example NZP2213, can induce organogenesis, but the nodules that form do not become infected (Pankhurst *et al*., 1979). There is even evidence for a defense response via flavolans produced by *L. pedunculatus.* It has also been shown that *L. pedunculatus* increases the production of certain flavonoids exclusively in non-fixing nodules (Cooper & Rao, 1992). The strain NZP2213 is sensitive to the plant produced prodelphinidin, which probably leads to the nod$^+$fix$^-$ phenotype (Pankhurst & Jones, 1979).

Chitin acts as a MAMP and induces PAMP-triggered immunity (PTI) in plants and animals (Nürnberger & Brunner, 2002). Plant receptor kinases involved in chitin signaling, including *At*-CERK1 or the co-receptor BRI1-ASSOCIATED KINSASE (BAK1), are targeted by the bacterial effector AvrPtoB from *Pseudomonas syringae* (Boller & Felix, 2009). AvrPtoB ubiquitinates and targets *At*-CERK1 for degradation and inhibits the functional hetero-dimerization of BAK1 with

other receptors (Boller & Felix, 2009; Chinchilla *et al.*, 2009; Gimenez-Ibanez *et al.*, 2009). NFs and chitin are structurally similar and chitin as well as NF signaling are linked to bacterial effector signaling. In *B. japonicum*, the expression of the type III gene cluster (*tts*) that encodes the T3SS is controlled in a similar way to NF production via flavonoid induced NodD activation (Krause *et al.*, 2002). The host targets of these bacterial effectors have not yet been identified. However, it is conceivable that these targets involved in rhizobia recognition are sensed in an analogous manner as pathogen triggered immunity. In our study, we identified the kinase domain of *NFR5* as one of two regions with signatures of positive selection. These regions may be targets for strain specific rhizobial signals. While it is likely that NFs bind to the ectodomains, specifically the LysM domains of NFR5 and NFR1, rhizobial type III effectors may target the NFR kinase domain by stabilizing or destabilizing receptor complexes. Binding of bacterial effectors with e3-ligase activity and removal of the receptors, or endocytotic signaling might be one possible scenario (Göhre *et al.*, 2008).

A point mutation within the LsyM2 domain of *At-CERK1*, the closest ortholog *of Lj-NFR1*, at a position homologous to position L133F of *Lj*-NFR5, leads to a deregulated defense response upon powdery mildew infections, where the fungus occupies the plant cells without cell death of the plant (Lipka, personal communication, October 16, 2008). A radical transformation from a non-host to a host has yet to be reported between legumes and rhizobia. Studies on the contribution to specificity of the LysM domains or ectodomains of NFR1 and NFR5 failed to show such a strong host switch in *Lotus* (Radutoiu *et al.*, 2007; Bek *et al.*, 2010).

V.2.3 Coevolutionary forces link selection and exclusion

Foster and Kokko (2006) suggested that mechanisms such as partner choice and cooperation would be unstable by reducing the variation in the symbiont. Presuming that a certain level of variability ensures genetic variation and thus persistence of the symbiosis, legumes must keep several doors open by preserving the different modes of bacterial entry correlated with varying specificity requirements towards rhizobia (Sprent, 2007). This would allow the plant to choose highly compatible rhizobia via the fast track root hair IT entry passage and would still allow a certain level of non-compatibility for a wider range of bacteria entering via cracks in the epidermis. However, by keeping the door open for less compatible bacteria, legume hosts face

the risk that rhizobial cheaters or pathogenic bacteria can enter the host cells. Given that bacterial horizontal gene transfer can be frequent in the rhizosphere, non-fixing bacteria can potentially acquire nodulation ability. Therefore, the legume host must have an independent surveillance system able to recognize non-symbiotic bacteria. Controlled exclusion could act at a deeper level of the RNS by discrimination against bacteria recognized as pathogens before release from the IT or entry into the plant cell from the intercellular space. We suggest that two mechanisms act on the RNS, selection and exclusion. Selection probably acts via NF signaling and exclusion via the pathogen recognition pathway as part of the plant immunity.

Why do rhizobia maintain the effectors, as they obviously delimit their host range? Rhizobia are competing with each other for nodulation of putative hosts. The rhizobial T3SS was shown to have positive or negative effects on nodulation depending on the host species and, thus, probably contributes to the competitiveness of the rhizobia for nodulation of specific hosts.

In our co-inoculation experiments the nodulation of both *Lotus* species was similar to the levels of nodulation upon inoculation with the compatible rhizobium. This shows that both *Lotus* species can cope with mixed rhizobium populations under experimental conditions. Sachs *et al.* (2010) also provide evidence for the ability of the legume host, *Lotus strigosus*, to constrain the infection and later proliferation of a native-occurring bradyrhizobial cheater. *L. uliginosus* engages in RNS with the slow growing bradyrhizobia. We showed that it blocks the infection by fast growing mesorhizoba early during infection, and it additionally controls rhizobial entry at the time of release into the host cell from the IT. The mode of action of the host control of RNS is still obscure and might involve various lines of defense.

V.3 A quick T90 reporter line based assay to test the symbiotic activity of *Lotus japonicus* Gifu nodulation and AM mutants

V.3.1 <u>Symbiotic GUS activity of single and double homozygous T90 lines with symbiotic mutant background</u>

The newly generated double homozygous lines, T90 x *symrk-10* and *nfr1-1* x T90, together with two previously generated lines, *castor-2* x T90 and *ccamk-2* x T90 (Kistner *et al.*, 2005), were tested for rhizobium induction of T90 GUS activity in roots. After the application of *M. loti* strain R7A, NFs of the same strain, or *Rl*-10.2.1 no GUS activity was detectable in the homozygous individuals of lines: T90 x *symrk-*

10, nfr1-1 x T90 and *ccamk-2* x T90, in contrast to the reporter line T90 and the heterozygous individuals, in which GUS staining was detected. These results demonstrate that genes *NFR1, SYMRK* and *CCAMK* act upstream of the T90 GUS construct and are required for the induction of symbiotic GUS activity in the lines T90 x *symrk-10, nfr1-1* x T90, and *ccamk-2* x T90. Congruent with earlier results from Kistner *et al.* (2005), no GUS activity was detected in line *castor-2* x T90 upon inoculation with various rhizobia, demonstrating that *CASTOR* also acts upstream of the T90 GUS construct.

In terms of GUS activity, the reporter line T90 was fully responsive to all rhizobial stimuli and showed strong GUS activity, and formation of ITs, primordia, and nodules. In contrast, the untransformed as well as the hairy roots of T90 lines with symbiotic mutant background transformed with the empty vector were unresponsive to *M. loti* and *Bradyrhizobium* sp. and only partially responsive to *Rl*-10.2.1 at an early time point post inoculation. Especially the symbiotic GUS capacity of line *nfr1-1* x T90 varied depending on the rhizobium strain used. Both strains *M. loti* and *Bradyrhizobium* sp. failed to induce GUS activity on line *nfr1-1* x T90, while strain *Rl*-10.2.1 induced weak GUS activity independent of *NFR1*. This supports the hypothesis of an alternative entry pathway independent of RHC and probably independent of *NFR1*. On the wild-type of *L. japonicus* Gifu the strain *Rl*-10.2.1 was tested nod⁻, but the strain induced the formation of bumps and elongated infection zones (EIZ) on roots of *L. japonicus* Nepal. A cross between the reporter line T90 and the wild-type *L. japonicus* Nepal could help identify the genes encoding this polymorphism via a segregation and gene mapping analysis.

In some double homozygous T90 plants with symbiotic mutant line background, the strains *M. loti* and *Bradyrhizobium* sp. induced GUS activity that was restricted to a single or a few rhizodermal cells. Probably these represent infection attempts that were aborted early but after the induction of GUS. On one individual of line T90 x *symrk-10,* a GUS⁺ stained bump formed upon inoculation with *M. loti*. This demonstrates a certain leakiness of the *symrk-10* nod⁻ phenotype.

V.3.2 Symbiotic GUS activity in hairy roots of double homozygous T90 lines with symbiotic mutant background expressing T265D gain-of-function construct of CCAMK

Hairy roots of all tested double homozygous T90 symbiosis mutant lines, *nfr1-1* x T90, T90 x *symrk-10, castor-2* x T90 and *ccamk-2* x T90, transformed with T265D gain-of-function construct of CCAMK displayed a rhizobium independent GUS activity. No GUS staining was detected in the hairy roots of the same lines transformed with the empty vector control. These results show that the symbiotic T90 GUS expression is activated via the common *SYM* pathway and can be induced by the autoactive version of CCAMK.

ITs and later infected nodules developed in T265D transformed roots of the T90 x *symrk-10* transgenic line inoculated with *M. loti*. This confirms the results by Madsen et al. (2010) showing that the infection deficiency of *symrk* mutants is rescued by T265D and that *CCAMK* is required for organogenesis. In the T265D transformed hairy roots of lines *nfr1-1* x T90, T90 x *symrk-10, castor-2* x T90 and *ccamk-2* x T90 GUS staining was detected in rhizodermal cells including root hairs.

V.3.3 Symbiotic and non-symbiotic GUS activity in vacuum infiltrated seedlings of double homozygous T90 lines with symbiotic mutant background

In contrast to the hairy roots of all tested T90 lines with symbiotic mutant background, which did not show any GUS staining upon rhizobium application, aberrant GUS activity was detected in some vacuum infiltrated plants under different conditions involving both rhizobia and the T265D construct. After mock-infiltration and mock-inoculation a faint GUS staining was visible in one plant each of the lines T90, *nfr1-1* x T90, T90 x *symrk-10*, and *ccamk-2* x T90.

This demonstrates the presence of a rhizobium and common *SYM* independent T90 GUS activity and suggests the existence of an alternative pathway, which leads to the activation of non-symbiotic GUS in the reporter line T90. Yet it is unclear how this GUS activity is triggered. It could relate to abiotic stress, e.g. the infiltration and co-cultivation media, the shifting from one medium to another, or the variable light conditions. Large areas of the roots were devoid of root hairs and revealed a heterogeneous growth phenotype with a proximal broad root diameter and a protuberance at the transition to a long and slender distal part of the root.

After infiltration with *A. rhizogenes* AR1193 carrying the empty vector control a moderate GUS staining was detected in all roots of reporter line T90 and in

some roots of lines *nfr1-1* x T90, T90 x *symrk-10*, and *castor-2* x T90. But no GUS staining was detected in line *ccamk-2* x T90. Generally the GUS staining was stronger in roots of all symbiosis mutant lines that were infiltrated with *A. rhizogenes* even when carrying the empty vector control. These results point to a *NFR1*, *SYMRK*, and *CASTOR* independent, but *CCAMK*-dependent pathway that activates symbiotic T90 GUS activity. Interestingly, NFR1, SYMRK and CASTOR act upstream of calcium spiking are not required to induce this GUS activity but *CCAMK* that is thought to decode the calcium spiking signal is required for rhizobium independent GUS activity in T90.

The T90 related GUS activity seems inducible also via a response pathway to biotic stress. It is easily conceivable that compounds with signaling functions in plant responses to abiotic and biotic stresses, such as jasmonic acid (JA), ethylene, salicylic acid (SA), or abscisic acid (ABA) induce the common *SYM* independent GUS activity. However, in the epidermis, defense hormones including JA, ethylene, and SA, and the stress hormone ABA negatively regulate NF induced calcium spiking (Ding & Oldroyd, 2009). Interestingly, in all tested roots GUS staining was not in epidermal cells but in the cell layers below. The T90 GUS insertion is located in the promoter of *CBP1*, which has general features of a calcium binding protein. Thus it is likely that the calcium spiking signal in roots elicits or contributes to the T90 GUS activity.

Infiltration with *A. rhizogenes* carrying T265D gain-of-function construct of CCAMK and subsequent inoculation with *M. loti* led to GUS staining in 100% of the tested roots of lines T90, *castor-2* x T90, and *ccamk-2* x T90 and in more than 50% of the roots of lines *nfr1-1* and *symrk-10*. These results demonstrate an additive effect and show that rhizobia and T265D are able to induce GUS activity to a level higher than the non-symbiotic and *A. rhizogenes* induced GUS activity. This also supports the idea of two independent T90 GUS activation pathways. However, more tests are needed to identify and describe the stimuli that are able to induce the aberrant T90 GUS activity. Different hormones, the Myc-factor, different symbiotic or pathogenic bacteria and fungi could give further information about the role of T90 GUS activity and *CBP1* in *L. japonicus*.

VI REFERENCES

Aguilar, O, Riva, O, Peltzer, E. 2004. Analysis of *Rhizobium etli* and of its symbiosis with wild *Phaseolus vulgaris* supports coevolution in centers of host diversification. *Proceedings of the National Academy of Sciences, USA* **101**(37): 13548–13553.

Altschul, SF, Madden, TL, Schäffer, AA, Jinghui Zhang, ZZ, Miller, W, Lipman, DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**(17): 3389–3402.

Ané, J-M, Kiss, GB, Riely, BK, Penmetsa, RV, Oldroyd, GED, Ayax, C, Lévy, J, Debellé, F, Baek, J-M, Kalo, P*, et al.* **2004.** *Medicago truncatula DMI1* required for bacterial and fungal symbioses in legumes. *Science* **303**: 1364-1367.

Ardourel, M, Demont, N, Debellé, F, Maillet, F, de Billy, F, Promé, J-C, Dénarié, J, Truchet, G. 1994. *Rhizobium meliloti* lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. *The Plant Cell* **6**(10): 1357-1374.

Arnold, K, Bordoli, L, Kopp, J, Schwede, T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**(2): 195–201.

Banba, M, Siddique, AB, Kouchi, H, Izui, K, Hata, S. 2001. *Lotus japonicus* forms early senescent root nodules with *Rhizobium etli*. *Molecular Plant-Microbe Interactions* **14**(2): 173-180.

Barcellos, FG, Menna, P, Batista, JSdS, Hungria, M. 2007. Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* Inoculant Strain to Indigenous Diazotrophs *Sinorhizobium* (*Ensifer*) *fredii* and *Bradyrhizobium elkanii* in a Brazilian Savannah Soil. *Applied and Environmental Microbiology* **73**(8): 2635–2643.

Bateman, A, Bycroft, M. 2000. The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD). *Journal of Molecular Biology* **299**(4): 1113-1119.

Bechtold, N, Ellis, J, Pelletier, Pelletier, G. 1993. In planta *Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *CR Acad. Sci. Paris, Life Sci.* **316**: 1194-1199.

Bek, AS, Sauer, J, Thygesen, MB, Duus, JØ, Petersen, BO, Thirup, S, James, E, Jensen, KJ, Stougaard, J, Radutoiu, S. 2010. Improved characterization of Nod factors and genetically based variation in LysM receptor domains identify amino acids expendable for Nod factor recognition in *Lotus* spp. *Molecular Plant-Microbe Interactions* **23**(1): 58-66.

Beringer, JE. 1974. R factor transfer in *Rhizobium leguminosarum. Journal of General Microbiology* **84**: 188-198.

Bernal, A, Pan, Q, Pollack, J, Rose, L, Kozik, A, Willits, N, Luo, Y, Guittet, M, Kochetkova, E, Michelmore, RW. 2005. Functional analysis of the plant disease resistance gene *Pto* using DNA shuffling. *Journal of Biological Chemistry* **280**(24): 23073–23083.

Beynon, J, Buchanan-Wollaston, A, Setchell, S. 1978. High frequency transfer of nodulating ability between strains and species of *Rhizobium. Nature* **276**: 634-636.

Boisson-Dernier, Al, Chabaud, M, Garcia, F, Bécard, G, Rosenberg, C, Barker, DG. 2001. *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Molecular Plant-Microbe Interactions* **4**(6): 695–700.

Boller, T, Felix, G. 2009. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology* **60**: 379-406.

Bonfante, P, Genre, A, Faccio, A, Martini, I, Schauser, L, Stougaard, J, Webb, J, Parniske, M. 2000. The *Lotus japonicus LjSym4* gene is required for the successful symbiotic infection of root epidermal cells. *Molecular Plant-Microbe Interactions* **13**(10): 1109-1120.

Bontemps, C, Elliott, G, Simon, M, Nior, FBdDJ, Gross, E, Lawton, RC, Neto, NE, Loureiro, MdF, Faria, SMd, Sprent, JI*, et al.* 2010. *Burkholderia* species are ancient symbionts of legumes. *Molecular Ecology* **19**: 44-52.

Broughton, W, Zhang, F, Perret, X, Staehelin, C. 2003. Signals exchanged between legumes and Rhizobium: agricultural uses and perspectives. *Plant and Soil* **252**: 129-137.

Broughton, WJ, Jabbouri, S, Perret, X. 2000. Keys to symbiotic harmony. *Journal of Bacteriology* **182**(20): 5641-5652.

Burdon, J, Gibson, A, Searle, S, Woods, M, Brockwell, J. 1999. Variation in the effectiveness of symbiotic associations between native rhizobia and temperate Australian *Acacia*: within-species interactions. *Journal of Applied Ecology* **36**(3): 398-408.

Camp, ROd, Streng, A, Mita, S, Cao, Q, Polone, E, Liu, W, Ammiraju, JS, Kudrna, D, Wing, R, Untergasser, A*, et al.* 2010. LysM-Type Mycorrhizal Receptor Recruited for Rhizobium Symbiosis in Nonlegume *Parasponia. Science* DOI: 10.1126

Campanoni, P, Sutter, JU, Davis, CS, Littlejohn, GR, Blatt, MR. 2007. A generalized method for transfecting root epidermis uncovers endosomal dynamics in Arabidopsis root hairs. *The Plant Journal* **51**(2): 322-330.

Chabaud, M, Larsonneau, C, Marmouget, C, Huguet, T. 1996. Transformation of barrel medic (*Medicago truncatula* Gaertn.) by *Agrobacterium tumefaciens*

and regeneration via somatic embryogenesis of transgenic plants with the *MtENOD*12 nodulin promoter fused to the *gus* reporter gene. *Plant Cell Reports* **15**(5): 305-310.

Charpentier, M, Bredemeier, R, Wanner, G, Takeda, N, Schleiff, E, Parniske, M. 2008. *Lotus japonicus* CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis. *The Plant Cell* **20**(12): 3467-3479.

Chen, W-M, de Faria, SM, Straliotto, R, Pitard, RM, Simoes-Araujo, JL, Chou, J-H, Chou, Y-J, Barrios, E, Prescott, AR, Elliott, GN*, et al.* **2005a.** Proof that *Burkholderia* strains form effective symbioses with legumes: a study of novel *Mimosa*-nodulating strains from South America. *Applied and Environmental Microbiology* **71**(11): 7461-7471.

Chen, W-M, James, EK, Chou, J-H, Sheu, S-Y, Yang, S-Z, Sprent, JI. 2005b. β-rhizobia from *Mimosa pigra*, a newly discovered invasive plant in Taiwan. *New Phytologist* **168**(3): 661-675.

Chen, W-M, Moulin, L, Bontemps, C, Vandamme, P, Béna, G, Boivin-Masson, C. 2003. Legume symbiotic nitrogen fixation by β-proteobacteria is widespread in nature. *Journal of Bacteriology* **185**(24): 7266-7272.

Cheng, H-P, Walker, GC. 1998. Succinoglycan production by *Rhizobium meliloti* is regulated through the ExoS-ChvI two-component regulatory system. *Journal of Bacteriology* **180**(1): 20-26.

Chinchilla, D, Shan, L, He, P, de Vries, S, Kemmerling, B. 2009. One for all: the receptor-associated kinase BAK1. *Trends in Plant Science* **14**(10): 535-541.

Chua, K, Pankhurst, C, MacDonald, P, Hopcroft, DH, Jarvis, BDW, Scott, DB. 1985. Isolation and characterization of transposon Tn5-induced symbiotic mutants of *Rhizobium loti*. *Journal of Bacteriology* **162**(1): 335-343.

Coenye, T, Vandamme, P. 2003. Diversity and significance of *Burkholderia* species occupying diverse ecological niches. *Environmental Microbiology* **5**(9): 719-729.

Cooper, JE, Rao, JR. 1992. Localized changes in flavonoid biosynthesis in roots of *Lotus pedunculatus* after infection by *Rhizobium loti*. *Plant Physiology* **100**: 444-450.

D'Haeze, W, Mergaert, P, Promé, J, Holsters, M. 2000. Nod factor requirements for efficient stem and root nodulation of the tropical legume *Sesbania rostrata*. *Journal of Biological Chemitry* **275**(21): 15676–15684.

Degtjareva, GV, Kramina, TE, Sokoloff, DD, Samigullin, TH, Sandral, G, Valiejo-Roman, CM. 2008. New data on nrITS phylogeny of *Lotus* (Leguminosae, Loteae). *Wulfenia* **15**: 35-49.

Demchenko, K, Winzer, T, Stougaard, J, Parniske, M, Pawlowski, K. 2004. Distinct roles of *Lotus japonicus SYMRK* and *SYM15* in root colonization and arbuscule formation. *New Phytologist* **163**(2): 381-392.

Den Herder, G, Parniske, M. 2009. The unbearable naivety of legumes in symbiosis. *Current Opinion in Plant Biology* **12**(4): 491-499.

Díaz, C, Grønlund, M, Schlaman, H, Spaink, H. 2005. Induction of hairy roots for symbiotic gene expression studies. In: Márquez, AJ, Stougaard, J, Udvardi, M, Parniske, M, Spaink, H, Saalbach, G, Webb, J, Chiurazzi, M, Márquez, AJ, eds. *Lotus japonicus Handbook*. The Netherlands: Springer, 261-277.

Díaz, P, Borsani, O, Monza, J. 2005. *Lotus*-related species and their agronomic importance. In: Márquez, AJ, Stougaard, J, Udvardi, M, Parniske, M, Spaink, H, Saalbach, G, Webb, J, Chiurazzi, M, Márquez, AJ, eds. *Lotus japonicus Handbook*. The Netherlands: Springer, 25-37.

Ding, Y, Oldroyd, GED. 2009. Positioning the nodule, the hormone dictum. *Plant Signaling & Behavior* **4**(2): 89-93.

Dong, J, Xiao, F, Fan, F, Gu, L, Cang, H, Martin, GB, Chai, J. 2009. Crystal Structure of the Complex between *Pseudomonas* Effector AvrPtoB and the Tomato Pto Kinase Reveals Both a Shared and a Unique Interface Compared with AvrPto-Pto. *The Plant Cell* **21**: 1846-1859.

Elliott, GN, Chen, W-M, Chou, J-H, Wang, H-C, Sheu, S-Y, Perin, L, Reis, VM, Moulin, L, Simon, MF, Bontemps, C, *et al.* 2007. *Burkholderia phymatum* is a highly effective nitrogen-fixing symbiont of *Mimosa* spp. and fixes nitrogen ex planta. *New Phytologist* **173**: 168-180.

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

Fåhraeus, G. 1957. The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. *Journal of General Microbiology* **16**: 374-381.

Foster, K, Kokko, H. 2006. Cheating can stabilize cooperation in mutualisms. *Proceedings of the Royal Society B* **273**: 2233-2239.

Foster, K, Wenseleers, T. 2006. A general model for the evolution of mutualisms. *Journal of Evolutionary Biology* **19**: 1283–1293.

Gage, D. 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiology and Molecular Biology Reviews* **68**(2): 280–300.

Garau, G, Yates, R, Deiana, P, Howieson, JG. 2009. Novel strains of nodulating *Burkholderia* have a role in nitrogen fixation with papilionoid herbaceous legumes adapted to acid, infertile soils. *Soil Biology and Biochemistry* **41**: 125-134.

Gimenez-Ibanez, S, Hann, DR, Ntoukakis, V, Petutschnig, E, Lipka, V, Rathjen, JP. 2009. AvrPtoB targets the LysM receptor kinase CERK1 to promote bacterial virulence on plants. *Current Biology* **19**(5): 423-429.

Göhre, V, Spallek, T, Häweker, H, Mersmann, S, Mentzel, T, Boller, T, Torres, Md, Mansfield, JW, Robatzek, S. 2008. Plant pattern-recognition receptor FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB. *Current Biology* **18**: 1824–1832.

Goormachtig, S, Capoen, W, James, EK, Holsters, M. 2004. Switch from intracellular to intercellular invasion during water stress-tolerant legume nodulation. *Proceedings of the National Academy of Sciences, USA* **101**(16): 6303-6308.

Gough, C. 2003. Rhizobium symbiosis: insight into nod factor receptors. *Current Biology* **13**: R973-R975.

Grant, W, Small, E. 1996. The origin of the *Lotus corniculatus* (Fabaceae) complex: a synthesis of diverse evidence. *Canadian Journal of Botany* **74**(7): 975–989.

Gresshoff, P. 1997. Classical and molecular genetics of the model legume *Lotus japonicus*. *Molecular Plant-Microbe Interactions* **10**(1): 59-68.

Groth, M, Takeda, N, Perry, J, Uchida, H, Dräxl, S, Brachmann, A, Sato, S, Tabata, S, Kawaguchi, M, Wang, TL, et al. 2010. *NENA*, a *Lotus japonicus* homolog of Sec13, is required for rhizodermal infection by arbuscular mycorrhiza fungi and rhizobia but dispensable for cortical for Cortical Endosymbiotic Development. *The Plant Cell* **22**: 2509–2526.

Guex, N, Peitsch, M. 1997. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis* **18**: 2114-2723.

Hahn, M, Studer, D. 1986. Competitiveness of a nif- *Bradyrhizobium japonicum* mutant against the wild-type strain. *FEMS Microbiology Letters* **33**(1): 143-148.

Han, TX, Wanga, ET, Han, LL, Chen, WF, Sui, XH, Chen, WX. 2008. Molecular diversity and phylogeny of rhizobia associated with wild legumes native to Xinjiang, China. *Systematic and Applied Microbiology* **31**: 287–301.

Handberg, K, Stougaard, J. 1992. *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. *The Plant Journal* **2**(4): 487-496.

Hartmann, A, Schmid, M, Tuinen, D, Berg, G. 2009. Plant-driven selection of microbes. *Plant and Soil* **321**: 235–257.

Horvath, B, Bachem, C, Schell, J, Kondorosi, A. 1987. Host-specific regulation of nodulation genes in *Rhizobium* is mediated by a plant-signal, interacting with the *nodD* gene product. *The EMBO Journal* **6**(4): 841-848.

Iizasa, Ei, Mitsutomi, M, Nagano, Y. 2010. Direct binding of a plant LysM receptor-like kinase, LysM RLK1/CERK1, to chitin *in vitro*. *Journal of Biological Chemistry* **285**(5): 2996-3004.

Imaizumi-Anraku, H, Takeda, N, Charpentier, M, Perry, J, Miwa, H, Umehara, Y, Kouchi, H, Murakami, Y, Mulder, L, Vickers, K, et al. 2005. Plastid

proteins crucial for symbiotic fungal and bacterial entry into plant roots. *Nature* **433**(7025): 527-531.

Irisarri, P, Milnitsky, F, Monza, J, Bedmar, EJ. 1996. Characterization of rhizobia nodulating *Lotus subbiflorus* from Uruguayan soils. *Plant and Soil* **180**: 39-47.

James, E, Crawford, R. 1998. Effect of oxygen availability on nitrogen fixation by two *Lotus* species under flooded conditions. *Journal of Experimental Botany* **49**(320): 599-609.

James, EK, Sprent, JI. 1999. Development of N2-fixing nodules on the wetland legume *Lotus uliginosus* exposed to conditions of flooding. *New Phytologist* **142**: 219-231.

Jukes, TH, Cantor, CR. 1969. Evolution of protein molecules. In: Munro, HN, ed. *Mammalian Protein Metabolism.* New York: Academic Press, 21-132.

Justin, SHFW, Amstrong, W. 1987. The anatomical characteristics of roots and plant response to soil flooding. *New Phytologist* **106**: 465–495.

Kaku, H, Nishizawa, Y, Ishii-Minami, N, Akimoto-Tomiyama, C, Dohmae, N, Takio, K, Minami, E, Shibuya, N. 2006. Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proceedings of the National Academy of Sciences, USA* **103**(29): 11086-11091.

Kambara, K, Ardissone, S, Kobayashi, H, Saad, MM, Schumpp, O, Broughton, WJ, Deakin, WJ. 2009. Rhizobia utilize pathogen-like effector proteins during symbiosis. *Molecular Microbiology* **71**(1): 92-106.

Kanamori, N, Madsen, LH, Radutoiu, S, Frantescu, M, Quistgaard, EMH, Miwa, H, Downie, JA, James, EK, Felle, HH, Haaning, LL*, et al.* 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, USA* **103**(2): 359.

Kiers, ET, Denison, RF. 2008. Sanctions, cooperation, and the stability of plant-rhizosphere mutualisms. *Annual Review of Ecology, Evolution, and Systematics* **39**: 215–236.

Kistner, C, Parniske, M. 2002. Evolution of signal transduction in intracellular symbiosis. *TRENDS in Plant Science* **7**(11): 511-518.

Kistner, C, Winzer, T, Pitzschke, A, Mulder, L, Sato, S, Kaneko, T, Tabata, S, Sandal, N, Stougaard, J, Webb, KJ*, et al.* 2005. Seven *Lotus japonicus* genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. *The Plant Cell* **17**(8): 2217-2229.

Krause, A, Doerfel, A, Göttfert, M. 2002. Mutational and transcriptional analysis of the type III secretion system of *Bradyrhizobium japonicum. Molecular Plant-Microbe Interactions* **15**(12): 1228-1235.

Kunishima, S, Inoue, C, Kamiya, T, Ozawa, K. 2001. Presence of *Propionibacterium acnes* in blood components. *Transfusion* **41**: 1126-1129.

Laguerre, G, Nour, SM, Macheret, V, Sanjuan, J, Drouin, P, Amarger, N. 2001. Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. *Microbiology* **147**: 981-993.

Laus, M, van Brussel, A. 2005. Exopolysaccharide structure is not a determinant of host-plant specificity in nodulation of *Vicia sativa* roots. *Molecular Plant-Microbe Interactions* **18**(11): 1123–1129. .

Lavin, M, Herendeen, PS, Wojciechowski, MF. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology* **54**(4): 575–594.

Lévy, J, Bres, C, Geurts, R, Chalhoub, B, Kulikova, O, Duc, G, Journet, E-P, Ané, J-M, Lauber, E, Bisseling, T*, et al.* 2004. A putative Ca2+ and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science* **303**: 1361-1364.

Lewis, G, Schrire, B, Mackinder, B, Lock, M. 2005. *Legumes of the World*. Kew: Royal Botanic Gardens.

Lewis-Henderson, W, Djordjevic, MA. 1991. A cultivar-specific interaction between *Rhizobium leguminosarum* bv. *trifolii* and subterranean clover is controlled by nodM, other bacterial cultivar specificity genes, and a single recessive host gene. *Journal of Bacteriology* **173**(9): 2791-2799.

Librado, P, Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**(11): 1451–1452.

Limpens, E, Franken, C, Smit, P, Willemse, J, Bisseling, T, Geurts, R. 2003. LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* **302**(5645): 630-633.

Lohmann, GV, Shimoda, Y, Nielsen, MW, Jørgensen, FG, Grossmann, C, Sandal, N, Sørensen, K, Thirup, S, Madsen, LH, Tabata, S*, et al.* 2010. Evolution and regulation of the *Lotus japonicus* LysM receptor gene family. *Molecular Plant-Microbe Interactions* **23**(4): 510-521.

Long, SR. 1996. Rhizobium Symbiosis: Nod Factors in Perspective. *The Plant Cell* **8**: 1885-1898.

López-Lara, I, Blok-Tip, L, Quinto, C, Garcia, M, Stacey, G, Bloemberg, G, Lamers, G, Lugtenberg, B, Thomas-Oates, J, Spaink, H. 1996. NodZ of *Bradyrhizobium* extends the nodulation host range of *Rhizobium* by adding a fucosyl residue to nodulation signals. *Molecular Microbiology* **21**(2): 397-408.

Maddison, DR, Maddison, WP. 2005. MacClade: Analysis of phylogeny and character evolution. Version 4.08. Sunderland, Massachusetts: Sinauer Associates.

Madsen, EB, Madsen, LH, Radutoiu, S, Olbryt, M, Rakwalska, M, Szczyglowski, K, Sato, S, Kaneko, T, Tabata, S, Sandal, N*, et al.* 2003. A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* **425**(6958): 637-640.

Madsen, LH, Tirichine, Ll, Jurkiewicz, A, Sullivan, JT, Heckmann, AB, Bek, AS, Ronson, CW, James, EK, Stougaard, J. 2010. The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nature communications* **1**(1): 1-12.

Maekawa, T, Kusakabe, M, Shimoda, Y, Sato, S, Tabata, S, Murooka, Y, Hayashi, M. 2008. Polyubiquitin Promoter-Based Binary Vectors for Overexpression and Gene Silencing in *Lotus japonicus. Molecular Plant-Microbe Interactions* **21**(4): 375-382.

Maillet, F, Poinsot, V, André, O, Puech-Pagès, V, Haouy, A, Gueunier, M, Cromer, L, Giraudet, D, Formey, D, Niebel, A*, et al.* 2011. Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. *Nature* **469**(7328): 58-63.

Marchetti, M, Capela, D, Glew, M, Cruveiller, S, Chane-Woon-Ming, B, Gris, C, Timmers, T, Poinsot, V, Gilbert, LB, Heeb, P*, et al.* 2010. Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biology* **8**(1): e1000280.

Markmann, K, Giczey, G, Parniske, M. 2008. Functional adaptation of a plant receptor-kinase paved the way for the evolution of intracellular root symbioses with bacteria. *PLoS Biol* **6**(3): 497–506.

Markmann, K, Parniske, M. 2009. Evolution of root endosymbiosis with bacteria: How novel are nodules? *Trends in Plant Science* **14**(2): 77-86.

Martiran, L, Stiller, J, Mirabella, R, Alfano, F, Lamberti, A, Radutoiu, SE, Iaccarino, M, Gresshoff, PM, Chiurazzi, M. 1999. T-DNA tagging of nodulation-and root-related genes in *Lotus japonicus*: Expression patterns and potential for promoter trapping and insertional mutagenesis. *Molecular Plant-Microbe Interactions* **12**(4): 275-285.

Mergaert, P, Uchiumi, T, Alunni, B, Evanno, G, Cheron, A, Catrice, O, Mausset, A-E, Barloy-Hubler, F, Galibert, F, Kondorosi, A*, et al.* 2006. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*–legume symbiosis. *Proceedings of the National Academy of Sciences, USA* **103**(13): 5230-5235.

Messinese, E, Mun, J-H, Yeun, LH, Jayaraman, D, Rougé, P, Barre, A, Lougnon, G, Schornack, S, Bono, J-J, Cook, DR*, et al.* 2007. A novel nuclear protein interacts with the symbiotic DMI3 calcium-and calmodulin-dependent protein kinase of *Medicago truncatula. Molecular Plant-Microbe Interactions* **20**(8): 912-921.

Mitra, R, Shaw, S, Long, S. 2004. Six nonnodulating plant mutants defective for Nod factor-induced transcriptional changes associated with the legume-

rhizobia symbiosis. *Proceedings of the National Academy of Sciences, USA* **101**(27): 10217-10222.

Miwa, H, Sun, J, Oldroyd, G, Downie, J. 2006. Analysis of calcium spiking using a cameleon calcium sensor reveals that nodulation gene expression is regulated by calcium spike number and the developmental status of the cell. *Plant J.* **48**(6): 883-894.

Miya, A, Albert, P, Shinya, T, Desaki, Y, Ichimura, K, Shirasu, K, Narusaka, Y, Kawakami, N, Kaku, H, Shibuya, N. 2007. CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proceedings of the National Academy of Sciences, USA* **104**(49): 19613-19618.

Moulin, L, Munive, A, Dreyfus, B, Boivin-Masson, C. 2001. Nodulation of legumes by members of the β-subclass of Proteobacteria. *Nature* **411**(6840): 948-950.

Mulder, L, Lefebvre, B, Cullimore, J, Imberty, A. 2006. LysM domains of *Medicago truncatula* NFP protein involved in Nod factor perception. Glycosylation state, molecular modeling and docking of chitooligosaccharides and Nod factors. *Glycobiology* **16**(9): 801-809.

Nawrath, SM. 2005. *Flora und Vegetation des Grünlands im südöstlichen Taunus und seinem Vorland.* J.W. Goethe-Universität Frankfurt am Main.

Nei, M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia Univ. Press.

Nei, M, Kumar, S. 2000. *Molecular evolution and phylogenetics*. New York: Oxford University Press, Inc.

Nielsen, R, Yang, Z. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* **148**(3): 929-936.

Nürnberger, T, Brunner, F. 2002. Innate immunity in plants and animals: emerging parallels between the recognition of general elicitors and pathogen-associated molecular patterns. *Current Opinion in Plant Biology* **5**: 1-7.

Okazaki, S, Okabe, S, Higashi, M, Shimoda, Y, Sato, S, Tabata, S, Hashiguchi, M, Akashi, R, Göttfert, M, Saeki, K. 2010. Identification and functional analysis of type III effector proteins in *Mesorhizobium loti. Molecular Plant-Microbe Interactions* **23**(2): 223-234.

Oldroyd, GED, Downie, JA. 2004. Calcium, Kinases and Nodulation Signalling in Legumes. *Nature Reviews: Molecular Cell Biology* **5**: 566-576.

Ooki, Y, Banba, M, Yano, K, Maruya, J, Sato, S, Tabata, S, Saeki, K, Hayashi, M, Kawaguchi, M, Izui, K. 2005. Characterization of the *Lotus japonicus* symbiotic mutant *lot1* that shows a reduced nodule number and distorted trichomes. *Plant Physiology* **137**(4): 1261-1271.

Ovtsyna, AO, Dolgikh, EA, Kilanova, AS, Tsyganov, VE, Borisov, AY, Tikhonovich, IA, Staehelin, C. 2005. Nod factors induce nod factor cleaving

enzymes in pea roots. Genetic and pharmacological approaches indicate different activation mechanisms. *Plant Physiology* **139**: 1051–1064.

Pacios Bras, C, Jorda, MA, Wijfjes, AHM, Harteveld, M, Stuurman, N, Thomas-Oates, JE, Spaink, HP. 2000. A *Lotus japonicus* nodulation system based on heterologous expression of the fucosyl transferase NodZ and the acetyl transferase NolL in *Rhizobium leguminosarum. Molecular Plant-Microbe Interactions* **13**(4): 475-479.

Pankhurst, CE, Craig, AS, Jones, WT. 1979. Effectiveness of *Lotus* root nodules: I. Morphology and flavolan content of nodules formed on *Lotus pedunculatus* by fast-growing *Lotus rhizobia. Journal of Experimental Botany* **30**: 1085-1093.

Pankhurst, CE, Jones, WT. 1979. Effectiveness of *Lotus* root nodules: II. Relationship between root nodule effectiveness and "in vitro" sensitivity of fast-growing *Lotus* rhizobia to flavolans. *Journal of Experimental Botany* **30**: 1095-1107.

Parniske, M. 2008. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nature Reviews Microbiology* **6**(10): 763-775.

Parniske, M, Ahlborn, B, Werner, D. 1991. Isoflavonoid-inducible resistance to the phytoalexin glyceollin in soybean rhizobia. *Journal of Bacteriology* **173**(11): 3432-3439.

Perret, X, Staehelin, C, Broughton, WJ. 2000. Molecular basis of symbiotic promiscuity. *Microbiology and Molecular Biology Reviews* **64**(1): 180-201.

Perry, J, Brachmann, A, Welham, T, Binder, A, Charpentier, M, Groth, M, Haage, K, Markmann, K, Wang, TL, Parniske, M. 2009. TILLING in *Lotus japonicus* identified large allelic series for symbiosis genes and revealed a bias in functionally defective ethyl methanesulfonate alleles toward glycine replacements. *Plant Physiology* **151**(3): 1281-1291.

Perry, J, Wang, T, Welham, T, Gardner, S, Pike, J, Yoshida, S, Parniske, M. 2003. A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus. Plant Physiology* **131**(3): 866-871.

Radutoiu, S, Madsen, LH, Madsen, EB, Felle, HH, Umehara, Y, Gronlund, M, Sato, S, Nakamura, Y, Tabata, S, Sandal, N, *et al.* 2003. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425**(6958): 585-592.

Radutoiu, S, Madsen, LH, Madsen, EB, Jurkiewicz, A, Fukai, E, Quistgaard, EMH, Albrektsen, AS, James, EK, Thirup, S, Stougaard, J. 2007. LysM domains mediate lipochitin–oligosaccharide recognition and *Nfr* genes extend the symbiotic host range. *The EMBO Journal* **26**(17): 3923-3935.

Rathjen, JP, Chang, JH, Staskawicz, BJ, Michelmore, RW. 1999. Constitutively active *Pto* induces a *Prf*-dependent hypersensitive response in the absence of *avrPto*. *The EMBO Journal* **18**(12): 3232–3240.

Rodpothong, P, Sullivan, JT, Songsrirote, K, Sumpton, D, Cheung, KWJ-T, Thomas-Oates, J, Radutoiu, S, Stougaard, J, Ronson, CW. 2009. Nodulation Gene Mutants of *Mesorhizobium loti* R7A-nodZ and nolL Mutants Have Host-Specific Phenotypes on *Lotus* spp. *Molecular Plant-Microbe Interactions* **22**(12): 1546-1554.

Röhm, M, Werner, D. 1991. Nitrate levels affect the development of the black locust-*Rhizobium* symbiosis. *Trees-Structure and Function* **5**: 227-231.

Sachs, J, Russell, J, Lii, Y, Black, KC, Lopez, G, Patil, AS. 2010. Host control over infection and proliferation of a cheater symbiont. *Journal of Evolutionary Biology* **23**(9): 1919-1927.

Saeki, K, Kouchi, H. 2000. The *Lotus* symbiont, *Mesorhizobium loti*: molecular genetic techniques and application. *Journal of Plant Research* **113**: 457-465.

Saito, K, Yoshikawa, M, Yano, K, Miwa, H, Uchida, H, Asamizu, E, Sato, S, Tabata, S, Imaizumi-Anraku, H, Umehara, Y, *et al.* 2007. NUCLEOPORIN85 Is Required for Calcium Spiking, Fungal and Bacterial Symbioses, and Seed Production in *Lotus japonicus*. *The Plant Cell* **19**(2): 610-624.

Sánchez, C, Iannino, F, Deakin, WJ, Ugalde, RA, Lepek, VC. 2009. Characterization of the *Mesorhizobium loti* MAFF303099 Type-Three Protein Secretion System. *Molecular Plant-Microbe Interactions* **22**(5): 519-528.

Sandal, N, Petersen, TR, Murray, J, Umehara, Y, Karas, B, Yano, K, Kumagai, H, Yoshikawa, M, Saito, K, Hayashi, M, *et al.* 2006. Genetics of symbiosis in Lotus japonicus: recombinant inbred lines, comparative genetic maps, and map position of 35 symbiotic loci. *Molecular Plant-Microbe Interactions* **19**(1): 80-91.

Schauser, L, Handberg, K, Sandal, N, Stiller, J, Thykjaer, T, Pajuelo, E, Nielsen, A, Stougaard, J. 1998. Symbiotic mutants deficient in nodule establishment identified after T-DNA transformation of Lotus japonicus. *Molecular and General Genetics MGG* **259**(4): 414-423.

Schumpp, O, Crèvecoeur, M, Broughton, WJ, Deakin, WJ. 2009. Delayed maturation of nodules reduces symbiotic effectiveness of the *Lotus japonicus-Rhizobium* sp. NGR234 interaction. *Journal of Experimental Botany* **60**(2): 581-590.

Scott, D, Young, CA, Collins-Emerson, JM, Terzaghi, EA, Rockman, ES, Lewis, PE, Pankhurst, CE. 1996. Novel and complex chromosomal arrangement of *Rhizobium loti* nodulation genes. *Molecular Plant-Microbe Interactions* **9**(3): 187-197.

Simon, R, Priefer, U, Pühler, A. 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Nature Biotechnology* **1**: 784-791.

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

Smit, P, Raedts, J, Portyanko, V, Debellé, F, Gough, C, Bisseling, T, Geurts, R. 2005. NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. *Science* **308**(5729): 1789-1791.

Somaroo, BH, Grant, WF. 1971. Interspecific hybridization between diploid species of *Lotus* (Leguminosae). *Genetica* **42**: 353-367.

Spaink, H, Wijffelman, C, Pees, E, Lugtenberg, ROBJJ. 1987. *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. *nature.com* **328**: 337-340.

Spaink, HP, Okker, RJH, Wijffelman, CA, Pees, E, Lugtenberg, BJJ. 1987. Promoters in the nodulation region of the Rhizobium leguminosarum Sym plasmid pRL1JI. *Plant Molecular Biology* **9**: 27-39.

Sprent, JI. 2007. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytologist* **174**(1): 11-25.

Sprent, JI. 2008. 60Ma of legume nodulation. What's new? What's changing? *Journal of Experimental Botany* **59**(5): 1081-1084.

Staden, R. 1996. The Staden sequence analysis package. *Molecular Biotechnology* **5**: 233-241.

Staehelin, C, Forsberg, LS, D'Haeze, W, Gao, M-Y, Carlson, RW, Xie, Z-P, Pellock, BJ, Jones, KM, Walker, GC, Streit, WR*, et al.* 2006. Exo-oligosaccharides of *Rhizobium* sp. strain NGR234 are required for symbiosis with various legumes. *Journal of Bacteriology* **188**(17): 6168-6178.

Stiller, J, Martirani, L, Tuppale, S, Chian, R, Chiurazzi, M, Gresshoff, P. 1997. High frequency transformation and regeneration of transgenic plants in the model legume *Lotus japonicus*. *Journal of Experimental Botany* **48**(7): 1357-1365.

Stougaard, J, Abildsten, D, Marcker, K. 1987. The *Agrobacterium rhizogenes* pRi TL-DNA segment as a gene vector system for transformation of plants. *Molecular and General Genetics MGG* **207**(2): 251-255.

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

Sullivan, J, Patrick, H, Lowther, W, Scott, D, Ronson, C. 1995. Nodulating Strains of *Rhizobium loti* Arise through Chromosomal Symbiotic Gene Transfer in the Environment. *Proceedings of the National Academy of Sciences, USA* **92**(19): 8985-8989.

Sullivan, JT, Ronson, CW. 1998. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proceedings of the National Academy of Sciences, USA* **95**(9): 5145-5149.

Sutton, JM, Lea, EJA, Downie, JA. 1994. The nodulation-signaling protein NodO from *Rhizobium leguminosarum* biovar *viciae* forms ion channels in membranes. *Proceedings of the National Academy of Sciences* **91**: 9990-9994.

Swanson, EB, Somers, DA, Tomes, DT. 1990. Birds foot treefoil (*Lotus corniculatus* L.) In: Bajaj, YPS, ed. *Biotechnology in Agriculture and Forestry*. Berlin: Springer-Verlag.

Swofford, D. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, Massachusetts: Sinauer Associates.

Thompson, J, Gibson, T, Plewniak, F, Jeanmougin, F, Higgins, D. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**(24): 4876–4882.

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

Vincent, JM. 1970. *A manual for the practical study of root nodule bacteria*. Oxford, United Kingdom: Blackwell Scientific.

Vinuesa, P, Rademaker, J, De Bruijn, F, Werner, D. 1998. Genotypic characterization of *Bradyrhizobium* strains nodulating endemic woody legumes of the Canary Islands by PCR-restriction fragment length polymorphism analysis of genes encoding 16S rRNA (16S rDNA) and 16S-23S rDNA intergenic spacers, repetitive extragenic palindromic PCR genomic fingerprinting, and partial 16S rDNA sequencing. *Applied and Environmental Microbiology* **64**(6): 2096–2104.

Walker, SA, Downie, JA. 2000. Entry of *Rhizobium leguminosarum* bv.*viciae* into root hairs requires minimal Nod factor specificity, but subsequent infection thread growth requires *nodO* or *nodE*. *Molecular Plant-Microbe Interactions* **13**(7): 754-762.

Wan, J, Zhang, X, Neece, D, Ramonell, K, Clough, S, Kim, S, Stacey, M, Stacey, G. 2008. A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in *Arabidopsis*. *The Plant Cell* **20**(2): 471-481.

Webb, KJ, Skøt, L, Nicholson, MN, Jørgensen, B, Mizen, S. 2000. *Mesorhizobium loti* increases root-specific expression of a calcium-binding protein homologue identified by promoter tagging in *Lotus japonicus*. *Molecular Plant-Microbe Interactions* **13**(6): 606-616.

Webster, G, Poulton, P, Cocking, E, Davey, M. 1995. The nodulation of micro-propagated plants of *Parasponia andersonii* by tropical legume rhizobia. *Journal of Experimental Botany* **46**(9): 1131-1137.

Wijffelman, CA, Zaat, B, Spaink, H, Mulders, I, Van Brussel, AAN, Okker, R, Pees, E, De Maagd, R, Lugtenberg, BJJ. 1986. Induction of *Rhizobium nod* genes by flavonoids: differential adaptation of promoter, *nodD* gene and inducers for various cross-inoculation groups. In: Lugtenberg, B, ed. *Recognition in Microbe-Plant Symbiotic and Pathogenic Interactions.* Heidelberg, Germany: Springer Verlag, 123-135.

Wu, A, Andriotis, V, Durrant, M, Rathjen, JP. 2004. A patch of surface-exposed residues mediates negative regulation of immune signaling by tomato Pto kinase. *The Plant Cell* **16**: 2809–2821.

Xing, W, Zou, Y, Liu, Q, Liu, J, Luo, X, Huang, Q, Chen, S, Zhu, L, Bi, R, Hao, Q, *et al.* 2007. The structural basis for activation of plant immunity by bacterial effector protein AvrPto. *Nature* **449**(7159): 243-247.

Yang, S, Tang, F, Gao, M, Krishnan, H, Zhu, H. 2010. *R* gene-controlled host specificity in the legume–rhizobia symbiosis. *Proceedings of the National Academy of Sciences* **107**(43): 18735–18740.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**(8): 1586-1591.

Yang, Z, Nielsen, R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* **19**(6): 908-917.

Yang, Z, Nielsen, R, Goldman, N, Krabbe Pedersen, A-M. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**(1): 431-449.

Yang, Z, Swanson, WJ. 2002. Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Molecular Biology and Evolution* **19**(1): 49-57.

Yano, K, Yoshida, S, Müller, J, Singh, S, Banba, M, Vickers, K, Markmann, K, White, C, Schuller, B, Sato, S, *et al.* 2008. CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proceedings of the National Academy of Sciences, USA* **105**(51): 20540-20545.

Zehner, S, Schober, G, Wenzel, M, Lang, K, Göttfert, M. 2008. Expression of the *Bradyrhizobium japonicum* type III secretion system in legume nodules and analysis of the associated tts box promoter. *Molecular Plant-Microbe Interactions* **21**(8): 1087-1093.

Zhu, H, Riely, BK, Burns, NJ, Ane, J-M. 2006. Tracing Nonlegume Orthologs of Legume Genes Required for Nodulation and Arbuscular Mycorrhizal Symbioses. *Genetics* **172**(4): 2491-2499.

VII ACKNOWLEDGEMENTS

I wish to express my gratitude to all those who gave me the possibility to complete this thesis.

In particular I would like to thank my supervisor Prof. Dr. M. Parniske for giving me the opportunity to undertake this research in his laboratory. Additionally, I would like to thank him for his continual advise and support throughout this work and for the excellent scientific training.

My grateful thanks goes to Prof. Dr. Laura Rose for all the support she has provided towards the completion of this work, for motivating discussions, and for her constant encouragement. In this context I would also like to thank Dr. W. J. Deakin.

Moreover, I would like to take this opportunity to thank all members of the examination committee for their interest in my work and for their time they devoted to this thesis.

I would like to thank Joana Bittencourt-Silvestre for the careful reviewing of this thesis.

I also would like to thank all members of the laboratory and the biology department for establishing a great atmosphere that facilitates the exchange of knowledge. In particular I would like to thank Astrid Stück, Sylvia Singh, Dr. Andreas Brachmann, Dr. Ralf Heerman, Markus Kador, Dr. Takaki Maekawa, and Andreas Binder, for their practical support in the lab, for the exchange of ideas and suggestions, and for their encouragement.

Additionally, I thank K.-H. Braun, S. Kattari, R. Verma and T. Winzer for help sampling wild *Lotus* populations. I thank G. Strobel for help with the bacterial isolation. I thank T. E. Kramina for assistance with the morphological classification of the *Lotus* species. I thank the following individuals for providing bacterial strains and plasmids: W. J. Deakin for the NGR234 strains, K. Saeki for MAFF303099 *nifH* mutant, A. Becker for *E. coli* helper strain S17-1 and the gfp and rfp plasmids, M. Göttfert for the DT3S mutant strain, and Sylvia Singh for the gain-of-function construct T265D. I thank M. Rebuffo and Makoto Hayashi for seed material of *L. pedunculatus*, *L. glaber,* and *L. japonicus* Nepal. I thank Simone Lange for help with the cloning and sequencing of the *NFR5* alleles. I thank M. Antolin Llovera and J. Lambert for practical support with propagation and DNA isolation of the in-vitro clones.

This work was supported by the EU-INCO project LOTASSA to MP.

VIII APPENDIX

VIII.1 List of publications

VIII.1.1 <u>Publications in peer-reviewed journals</u>

Didier Salmon, Carsten Krumbholz, Markus Kador, Sabine Bachmaier, JASMIN A. GOSSMANN, Pierrick Uzureau, Etienne Pays, and Michael Boshart (submitted 17-Dec-2010): The ESAG4 sub-family of adenylyl cyclases is required for for cytokinesis of *Trypanosoma brucei* bloodstream forms. Submitted to *Molecular Microbiology*.

GOSSMANN, J. A., K. Markmann, L. E. Rose, M. Parniske (submitted 21-Sep-2010): Distribution of genetic variation and host specificity among nitrogen fixing symbionts from closely related *Lotus* species. Submitted to *New Phytologist*.

GOSSMANN, J. A., M. Parniske, M. Rebuffo, L. E. Rose (in preparation): Positively selected sequence signatures within the LysM 2 and kinase domains of Nod Factor Receptor 5 correlate with contrasting compatibilities of European populations of *Lotus pedunculatus* and *Lotus corniculatus* towards rhizobial symbionts. To be submitted to *New Phytologist*.

GOSSMANN, J. A., Piepenbring, M., Guaglianone, E. R. (2007): Smut fungi on species of *Rhynchospora* (Cyperaceae) from Argentina. *Nova Hedwigia*. 84(1-2): 2009-226

Matheny, P. B., GOSSMANN, J. A., Zalar, P., Kumar, T. K. A., and Hibbett, D. S. (2006): Resolving the phylogenetic position of the Wallemiomycetes: an enigmatic major lineage of Basidiomycota. *Canadian Journal of Botany* 84(12): 1794-1805

VIII.1.2 <u>Posters and conferences</u>

JASMIN A. GOSSMANN, Laura Rose and Martin Parniske (2008) "Loci determining specificity at steps post nodule induction in *Lotus*"; IIIème Cycle conference: Plants and microbes, from the cellular to the global level; 16th and 17th of October 2008 in Fribourg, Switzerland (POSTER)

JASMIN A. GOSSMANN, Laura Rose and Martin Parniske (2008) "Loci determining specificity at steps post nodule induction in *Lotus*"; 8th Europ. Nitrogen Fixation Conference; 30th of August – 3rd of September 2008 in Gent, Belgium (POSTER)

VIII.1.3 Talks

JASMIN A. GOSSMANN, Martin Parniske (2007) "Signal transduction in plant-microbe interactions"; International Workshop – Integrating genomics into plant breeding; 25th – 26th of October 2007 in Porto Alegre, Brazil (TALK)

JASMIN A. GOSSMANN, Laura Rose and Martin Parniske (2007) "Specificity conferring loci in *Lotus* sp."; 3rd Contractors' meeting: UFRGS; 22th – 24th of October 2007 in Porto Alegre, Brazil (TALK, PROGRESS REPORT)

JASMIN A. GOSSMANN, P. Brandon Matheny, David S. Hibbett, and Meike Piepenbring (2006) "*Entorrhiza* and the smut fungi (Ustilaginomycotina)" International Congress of the German Mycological Society; 29th of September – 7th of October 2006 in Tübingen, Deutschland (TALK)

JASMIN A. GOSSMANN, Meike Piepenbring (2005) "Molecular phylogenetics and systematics"; International mycological workshop in Chiriquí, Panama; 12th – 30th of September 2005 (TALK)

VIII.2 LIST OF FIGURES

VIII.3 LIST OF TABLES

VIII.4 Ehrenwörtliche Versicherung und Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorgelegte Dissertation über die

Multiple molecular components contribute to genotype specific compatibility of the root nodule symbiosis

selbständig angefertigt und mich nicht anderer Hilfsmittel als denen in ihr angegebenen bedient habe, insbesondere, dass aus Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche mit Angabe der betreffenden Schrift bezeichnet sind, nicht stattgefunden haben.

Ich habe weder anderweitig versucht, eine Dissertation einzureichen oder mich Doktorprüfung zu unterziehen, noch habe ich diese Dissertation oder Teile derselben einer anderen Prüfungskommission vorgelegt.

München, den................................   ....

Jasmin A. Gossmann

VIII.5 Curriculum vitae

EDUCATION:

PhD, Plant genetics, Ludwig-Maximilians-University in Munich, Germany, January 2011 (thesis submission)
Concentration: plant genetics and plant-microbe interactions
Dissertation topic: Multiple molecular components contribute to genotype specific
            compatibility of the root nodule symbiosis
Erstes Staatsexamen (Lehramt an Gymnasien), Biology and Sports, Johann Wolfgang Goethe-University of Frankfurt, Germany, May 2005
Concentration: plant biology, mycology, bioinformatics
Thesis: Morphological and taxonomic analysis of Argentinean species of the
         Cintractia-group (smut fungi) on *Rhynchospora* sp. (Cyperaceae).

FELLOWSHIPS

Visiting Student Scholar, July 2005 – February 2006
Department of Biology at Clark University, Worcester, MA, USA
Participation into the AFTOL (Assembling the Fungal Tree of Life) project

TEACHING EXPERIENCE

Supervision of diploma and undergraduate students
and teaching assistant and instructor, 2006-2009
Ludwig-Maximilians-University in Munich, Germany
Course: Bioinformatics (alignments and phylogenetic analyses), Genetics class

Student teaching assistant (forest, pasture, botanical garden guide), 2003-2004
Johann Wolfgang Goethe-University of Frankfurt, Germany
Course: Botany, Systematics

Student teaching assistant (tutor), 2003
Johann Wolfgang Goethe-University of Frankfurt, Germany
Course: Botany, Systematics

Student teaching assistant (tutor), 2002
Johann Wolfgang Goethe-University of Frankfurt, Germany
Course: General Botany and Zoology for undergraduates

PARTICIPATION IN SOCIAL ACTIVITIES

Mentor for girls and women in mathematics, informatics, life sciences, and engineering (MINT), 2010
Soccer trainer for girls, 2003-2004

LANGUAGES

Fluent in English and German
Proficiency in Spanish, Croatian, French