Genetic determination of symbiotic compatibility between *Lotus* species and rhizobia strains

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften
Doctor rerum naturalium (Dr. rer. nat.) an der Fakultät für Biologie der
Ludwig-Maximilians-Universität München

vorgelegt von

Yen-Yu Lin

München, March 2022

Diese Dissertation wurde angefertigt
unter der Leitung von PD Dr. Macarena Marín
im Bereich von Genetik
an der Ludwig-Maximilians-Universität München

Ersgutachter/in: PD Dr. Macarena Marín

Zweitgutachter/in: Prof. Dr. Silke Robatzek

Tag der Abgabe: 17 March 2022

Tag der mündlichen Prüfung: 22 June 2022

EIDESSTATTLICHE RKLÄRUNG

Ich versichere hier mit an Eides statt, dass meine Dissertation selbständig und ohne unerlaubte Hilfsmittel angefertigt worden ist.
Die vorliegende Dissertation wurde weder ganz, noch teilweise bei einer anderen Prüfungs kommission vorgelegt.
Ich habe noch zu keinem früheren Zeitpunkt versucht, eine Dissertation einzureichen oder an einer Doktorprüfung teilzunehmen.

München, den 29.06.2022

Yen-Yu Lin

Content

Abbreviations

°C	degree Celsius
µ	Micro
AAO	Aldehyde oxidase
ABA	Abscisic acid
ACC	1-aminocyclopropane-1-carboxylic acid
ANOVA	Analysis of variance
AON	Autoregulation of nodulation
APS	Adenosine 5′-phosphosulfate
ASL	*Asymmetric leaves 2-Like*
Au	*Allorhizobium undicola*
AVG	Aminoethoxyvinyl glycine
bp	Base pair
bv.	Biovar
CCaMK	Calcium calmodulin dependent kinase
CKX3	Cytokinin dehydrogenase 3
CLE	Clavata3/embryo surrounding region-related
COSY	Coumarin synthase
CRE1	Cytokinin Response 1
CRISPR	Clustered regularly interspaced short palindromic repeats
cv.	Cultivar
DEG	Differentially expressed genes
dpi	Days post ioculation
*Ds*Red	*Discosoma sp.* red fluorescent protein
ENOD	Early nodulin
Ef	*Ensifer fredii*
Ein2	Ethylene Insensitive 2
EPR3	Exopolysaccharide receptor 3
EPS	Exopolysaccharide
ERN1	Ethylene responsive factor required for nodulation 1
ETI	Effector-triggered immunity
F6’H1	Feruloyl-CoA 6′-Hydroxylase 1
GAs	Gibberellins
gDNA	Genomic DNA
GFP	Green fluorescent protein

GlcNAc	β-1,4-linked N-acetyl glucosamine
gRNA	Guide RNA
GUS	β-glucuronidase
GWAS	Genome-wide association studies
HM	High substrate moisture
h	hours
IPT	Isopentenyl transferase
IT	Infection thread
La	*Lachnospiraceae* sp.
LB	Lysogeny broth
Lb	*Lotus burttii*
LBD	*Lateral organ boundaries domain*
LHK	*Lotus* histidine kinase
Lj	*Lotus japonicus*
LM	Low substrate moisture
LPS	Lipopolysaccharides
LRR	Leucin rich repeats
LYK3	Lysin motif receptor-like kinase 3
LysM	Lysin motif
M	Molar
m	Milli
MAMP	Microbe-associated molecular pattern
MCSU	Molybdenum cofactor sulfurase
min	Minutes
miR	MicroRNA
Ml	*Mesorhizobium loti*
Mt	*Medicago truncatula*
MTI	MAMP-triggered immunity
NFR1	Nod factor receptor 1
NFR5	Nod factor receptor 5
n	Nano
NCED	9-cis-epoxycarotenoid dioxygenase
NF-Y	Nuclear factor Y
NIN	Nodule inception
Nod	Nodulation
NSP	Nodulation signaling pathway

PAL	Phenylalanine ammonia lyase
PAM	Protospacer-adjacent motif
PAMP	Pathogen-associated molecular pattern
PAPS	3′-phosphoadenosine 5′-phosphosulfate
PCR	Polymerase chain reaction
Ps	*Pisum sativum*
PTI	PAMP-triggered immunity
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
RBOH	Respiratory burst oxidase homologue
RIC	Rhizobia induced CLE
Rl	*Rhizobium leguminosarum*
RLK	Receptor-like kinase
RLP	Receptor-like protein
RNAi	RNA interference
ROS	Reactive oxygen species
RS	Root signal
SAR	Systemic acquired resistance
SYMREM1	Symbiotic remorin 1
SYMRK	Symbiosis receptor-like kinase
T3SS	Type 3 secretion system
T4SS	Type 4 secretion system
T6SS	Type 6 secretion system
TML	Too Much Love
ttCas12a	Temperature-tolerant Cas12a
TY	Tryptone yeast
QTL	Quantitative trait locus
s	Seconds
STS	Silver thiosulfate
tt	Temperature-tolerant
UMIs	Unique molecular identifiers
wpi	Weeks post-inoculation
YEB	Yeast extract broth
ZEP	Zeaxanthin epoxidase

List of Publications

Publication I:

Liang J**,** Klingl A**, Lin Y-Y**, Boul E, Thomas-Oates J, Marín M. (2019). A subcompatible *Rhizobium* strain reveals infection duality in *Lotus*. *J Exp Bot*, 70, 1903-1913.

Manuscript I:

Zarrabian M, Montiel J, Sandal N, Jin H, **Lin Y-Y**, Marín M, James E, Parniske M, Stougaard J and Andersen S U. A promiscuity locus confers *Lotus burtti* nodulation with rhizobia from five different genera. *bioRxiv*, doi: 10.1101/2021.08.26.457880

Manuscript II:

Lin Y-Y, Yu F-Y, Wange L, Enard W, Marín M. Substrate moisture modulates the symbiotic compatibility between sub-compatible partners. (In preparation)

Manuscript III:

Lin Y-Y, Grillmayer D, Malkaj E, Andersen S U, Parniske M, Marín M. Receptor-like protein-encoding genes contribute to the determination of symbiotic compatibility between *Lotus* species and rhizobia strains. (In preparation)

Summary

Legume crops greatly reduce the need for synthetic nitrogen fertilizers and thus have been indicated as central for sustainable agricultural practices. This results from a symbiosis with nitrogen-fixing rhizobia that provide legumes with nitrogen in exchange for carbohydrates in root organs known as nodules. In nature, legumes encounter the dilemma of whether to be selective on rhizobia symbiont but risk starvation or broaden the selectivity but increase the chance to be infected by ineffective rhizobia, which fix little to no nitrogen. Symbiosis between legumes and rhizobia is initiated after flavonoids in the root exudates induce the production of rhizobial Nod factors, which are perceived by the legume receptor complex. This molecular interaction determines the symbiotic compatibility between host and rhizobia species and triggers downstream rhizobia infection and root nodule organogenesis.

Variation in the symbiotic compatibility between legumes and rhizobia is observed between and within species. Here we investigate a *Rhizobium leguminosarum* strain Norway (*Rl* Norway) that nodulates different *Lotus* accessions without nitrogen fixation. Thus, it is considered a sub-compatible symbiont of *Lotus.* The most striking phenotype difference is between *L. burttii* and *L. japonicus* Gifu, as *Rl* Norway induces white nodules on *L. burttii* but fails to nodulate *L. japonicus* Gifu. A region associated with this variation in nodulation phenotype between *L. burttii* and *L. japonicus* Gifu was identified by quantitative trait locus (QTL) mapping, but the gene(s) responsible for the various symbiotic compatibility remained unknown. This study aimed to characterize candidate genes in the QTL region and reveal the putative regulatory mechanism mediating the symbiotic compatibility. To achieve this, phenotypic observation, transcriptomic sequencing, and genetic analysis were integrated.

We observed variation in symbiotic compatibility between *Lotus* accessions and *Ensifer* and *Allorhizobium* strains in addition to *Rl* Norway. Moreover, the substrate moisture affected the symbiotic compatibility between *L. japonicus* Gifu and *Rl* Norway, resulting in the nodulation of *L. japonicus* Gifu in high substrate moisture. Transcriptome analyses revealed that several genes involved in the flavonoid biosynthesis were downregulated in *L. japonicus* Gifu grown in high moisture. We hypothesized that the accumulation of intermediates in the flavonoid biosynthesis pathway alters the symbiotic compatibility in high moisture. Naringenin, a flavonoid compound that is predicted to accumulate in high moisture was applied to roots and its effect on nodule formation and rhizobia colonization was evaluated. Transcriptomic analyses also showed that *Rl* Norway activated the symbiosis response in *L. burttii* but not in *L. japonicus* Gifu. Phenotyping the F1 progeny of a cross between *L. burttii* and *L. japonicus*

Gifu indicated that the nodulation phenotype of *L. burttii* is dominant. Four genes encoding receptor-like proteins (RLPs) in the QTL region were identified as candidates that contribute to the symbiotic compatibility, named *RLP1* to *RLP4*. An additive effect between *RLP2* and *RLP4* on nodule formation was observed by trans-complementing *RLP2* and *RLP4* from *L. burttii* into *L. japonicus* Gifu. Furthermore, mutant lines of the *RLP*s were generated by the CRISPR-Cas12a gene-editing method for future studies. Altogether, we identified candidate genes that contribute to the different symbiotic compatibility between *L. burttii* and *L. japonicus* Gifu and showed the variation in symbiotic compatibility between *L. japonicus* Gifu and *Rl* Norway is altered by substrate moisture.

Introduction

1 Synthetic nitrogen fertilizers

Nitrogen is an indispensable nutrient for plant growth and reproduction. It is an essential component of amino acids, proteins, nucleic acids, chlorophylls, phytohormones, and some primary and secondary metabolites (Frink et al., 1999, O'Brien et al., 2016, Kishorekumar et al., 2019). Nitrate (NO_3^-) and ammonium (NH_4^+) are inorganic nitrogen forms that plants can absorb directly from the soil (Kishorekumar et al., 2019). Ammonium is often converted into nitrate via nitrification by soil microorganisms (Beeckman et al., 2018). However, nitrate is hard to preserve in soil due to its poor ability to form complexes with soil minerals (Beeckman et al., 2018). This makes nitrate easy to lose by leaching and denitrification, which causes up to a 50% decrease in nitrogen availability for plants (Garwood & Ryden, 1986, Dechorgnat et al., 2010, Beeckman et al., 2018). As the human population grows continuously, synthetic nitrogen fertilizers have become critical to replenish the nitrogen deficiency for high crop yields to feed the expected 9.8 billion people by 2050 (Mulvaney et al., 2009, Rigby et al., 2016, Dimkpa et al., 2020).

Synthetic nitrogen fertilizers are produced industrially by the Haber-Bosch process. This process was developed in the first decade of the 20 century and is one of the most significant contributions to public welfare (Jensen et al., 2011, Chen et al., 2019, Qing et al., 2020). In the Haber-Bosch process, ammonia is synthesized from nitrogen and hydrogen catalysis under energy-demanding conditions (500°C and 150-200 atm), which consume 2% of the global energy supply on average (Mulvaney et al., 2009, Foyer et al., 2016, Liu et al., 2020, Qing et al., 2020).

Nitrogen fertilizers are costly, and their production and excessive usage have resulted in environmental hazards (Smith, 2003, Mulvaney et al., 2009, Stevens, 2019). Greenhouse gases, such as carbon dioxide (CO_2) and nitrous oxide (N_2O), are emitted during the production of synthetic nitrogen fertilizers or via microbial-mediated nitrification in the soil after application (Hakeem et al., 2017, Stevens, 2019, Ahmed et al., 2020). The level of N_2O has elevated by 16% since 1750 at an annual increasing rate of 0.25% on average (Hakeem et al., 2017). Although N_2O represents a small portion of the atmosphere, it contributes 300 times more than CO_2 to global warming and accounts for 6% of the global warming process (Ahmed et al., 2020). In addition, the excess nitrogen leads to expanding eutrophication in water and further affects the ecosystems (Smith, 2003, Hakeem et al., 2017, Insausti et al., 2020). Therefore, sustainable agriculture practices have been the focus of research. Leguminous plants, the

dominant species in nitrogen-limiting environments (Zahran, 1999), are at the center of sustainable agriculture. They significantly reduce the need of nitrogen fertilizers since they can establish symbiosis with nitrogen-fixing rhizobia (Beltran et al., 2018).

2 Introduction to root nodule symbiosis

2.1 Legumes are at the center of sustainable agriculture

Leguminous plants are the third-largest angiosperm family and play important roles in ecological and agricultural systems (Sprent et al., 2017, Varshney et al., 2018). Seeds of legumes are nutritious in proteins and contain several kinds of vitamins and lipids (Verma et al., 2013, Beltran & Canas, 2018). For instance, chickpeas, lentils, beans, peas, lupins, and soybeans contain double to triple protein content than cereals (Wallace et al., 2016, Beltran & Canas, 2018). Therefore, legumes have been cultivated along with cereal crops and have become significant staples of proteins (Beltran & Canas, 2018).

In addition to their roles in food sustenance, leguminous plants greatly reduce the need for nitrogen fertilizers based on their ability to utilize atmospheric nitrogen by hosting symbiotic nitrogen-fixing rhizobia in specialized organs called nodules. In nodules, nitrogen-fixing rhizobia provide nitrogen to plants in exchange for carbohydrates provided by the hosts (Bhattacharyya & Jha, 2012, Foyer et al., 2019). This symbiosis between legumes and rhizobia can contribute up to 80% of the biological nitrogen fixation in agricultural practices (Mabrouk et al., 2018) and reduce the usage of nitrogen fertilizers in cropping systems (Lengwati et al., 2020).

The interaction between rhizobia and host plants could lead to adverse outcomes. While optimal partners can reach effective nitrogen fixation, plants with less selectivity form ineffective symbiosis with sub-compatible partners. In these associations, rhizobia fix little to no nitrogen for plants, which leads to reduced crop production (Terpolilli et al., 2008, Friesen, 2012, Israel et al., 1986). Thus, breeding plants with higher selectivity is agronomically relevant.

2.2 Overview of root nodule symbiosis

Via an intricate molecular cross-talk, legumes and rhizobia establish a mutualistic root nodule symbiosis when the compatible partners are matched. In the beginning, rhizobia are attracted by plant exudates in the rhizosphere to colonize the roots. Flavonoids in the plant exudates activate rhizobial genes required during the symbiosis. After rhizobia perceive the flavonoids, the biosynthesis of the nodulation factors (Nod factors) is initiated (Recourt et al., 1989,

D'Haeze & Holsters, 2002). The Nod factors are then perceived by the Nod factor receptor complex (NFR1 and NFR5) on the host plasma membrane. This Nod factor perception initiates two concurrent processes, rhizobia infection and nodule organogenesis (Figure 1) (Oldroyd et al., 2011, Roy et al., 2020). In addition to the Nod factors, rhizobial exopolysaccharides and effectors also contribute to the host-rhizobia recognition process.

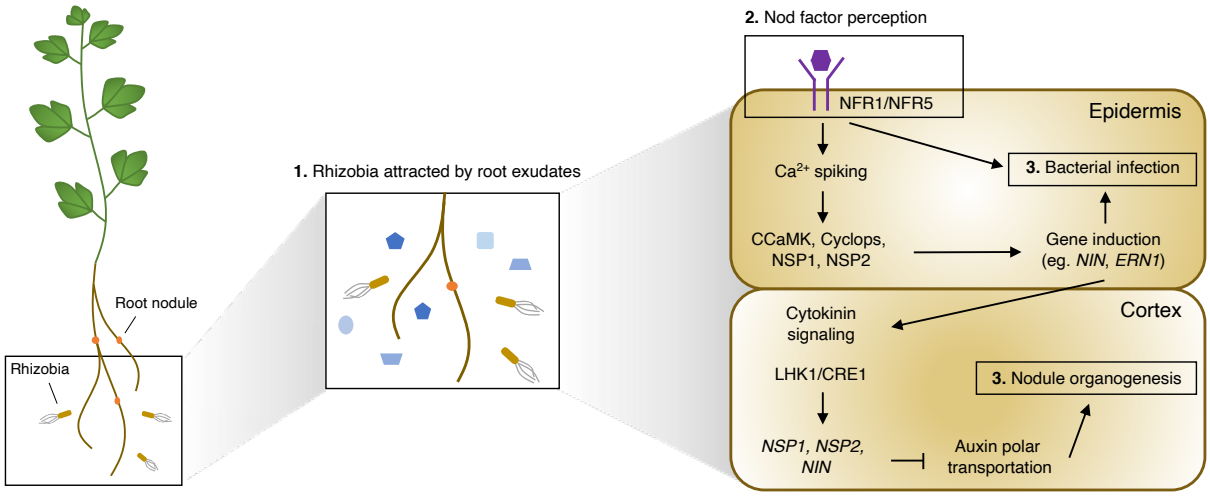


Figure 1. Illustration of steps in root nodule symbiosis establishment. In the rhizosphere, leguminous plants secrete exudates (depicted in different colors) to attract rhizobia (1). In response to the flavonoids, rhizobia produce Nod factors. The Nod factors are then perceived by the NFR1/NFR5 receptor complex on the membrane of root epidermal cells (2) and trigger downstream calcium oscillations and the simultaneous signaling pathways that lead to bacterial infection and nodule organogenesis (3). Scheme modified from Oldroyd et al., 2011.

After the molecular signal exchange between the host and compatible rhizobia, the rhizobia infection process begins at the root epidermis. To successfully infect the host plants, rhizobia need to cross the epidermis physically, spread in cortical cells, and be uptaken intracellularly by plant cells (Venado et al., 2020). Three infection strategies are categorized by how rhizobia overcome the epidermal barrier and spread in the cortex. The infection strategies include an infection thread- (IT) dependent infection via root hairs (Figure 2A) and two IT-independent mechanisms (Figure 2B and C). One of them is crack-entry infection via wounds and natural openings. Another mechanism is intercellular entry through middle lamella between epidermal cells (Ibanez et al., 2017, Sprent et al., 2017, Venado et al., 2020). Although evolutionary studies suggest the IT-independent infection emerges earlier, the IT-dependent infection has been broadly observed and studied in *Mimosoide*, *Cassieae*, *Caesalpinieae*, and *Papilionoid* clades, including model organisms *Lotus japonicus*, *Medicago truncatula*, and *Pisum sativum* (Sprent & James, 2007, Madsen et al., 2010).

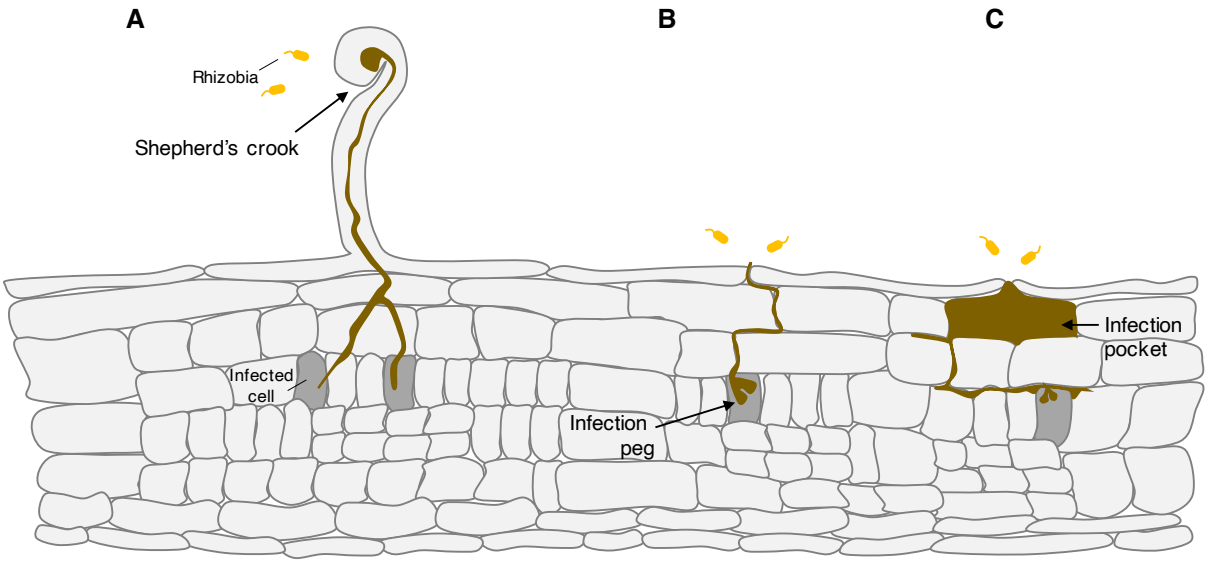


Figure 2. Rhizobia infection mechanisms. (A) Infection thread-dependent root hair infection. Rhizobia are entrapped at 'shepherd's crook' in curled root hairs. They reach the dividing cortical cells through the infection threads (ITs), then are internalized. (B) Intercellular infection. Rhizobia cross the epidermis between the middle lamella of adjacent cells, enter the cortical cells via 'infection pegs' and are then internalized. (C) Crack-entry infection. Rhizobia penetrate the root through fissures on the epidermis, for example, wounds or where lateral roots emerge. They accumulate in an infection pocket, from which they spread and infect other cells. This scheme is based on Venado et al., 2020.

3 Symbiotic perception and selectivity

3.1 Nod factor structure and Nod factor perception by the hosts

Nod factors contribute to host specificity and trigger downstream singling pathways that control bacterial infection and nodule organogenesis (Fliegmann, 2015; Oldroyd, 2011). They are lipochitooligosaccharides composed of an oligomeric backbone of β-1,4-linked *N*-acetyl-D-glucosamine (GlcNAc) units, an N-acyl group at the non-reducing terminal, which varies in length and degree of unsaturation of the fatty acyl group, and different chemical substituents at the reducing and non-reducing ends of the chitooligosaccharide backbone (D'Haeze, 2002; Fliegmann, 2015). These variations contribute to the host determination (Tikhonovich & Provorov, 2007).

The biosynthesis of the Nod factors starts from the activation of the *nodD* gene, which encodes a transcription factor responsive to flavonoids and initiates the transcription of *nod* operons for Nod factor production via binding to a conserved DNA sequence, the *nod* box (Wijffelman et al., 1989, Folch-Mallol et al., 1998). The lipooligosaccharide core of the Nod factors is synthesized by enzymes encoded by the *nodABC* genes. In brief, the GlcNAc is added to the non-reducing end of the chitin oligomer backbone by an *N*-acetyl-glucosaminyltransferase

encoded by *nodC* (Geremia et al., 1994, Kamst et al., 1997). The deacetylase encoded by *nodB* removes the N-acetyl group from the non-reducing end of the GlcNAc chitooligosaccharide (John et al., 1993). The acyltransferase encoded by *nodA* then attaches the fatty acid to the *N*-deacetylated end (Debelle et al., 1996, Tikhonovich & Provorov, 2007). This synthesized core of Nod factors is ready for substitution attachments at different ends.

Substitutions of Nod factors, such as acetyl, methyl, carbamoyl, and arabinosyl groups, are mediated by *nod*, *noe* and *nol* genes (Downie, 1998), which are rhizobia strain-specific and contribute to host-range determination (D'Haeze & Holsters, 2002). For instance, NolL of *Rhizobium* sp. strain NGR234 acetylates fucosylated residue of the Nod factors. Introducing the *nolL* gene of NGR234 strain into *Ensifer fredii* USDA257 extends the host range of this strain to *Calopogonium caeruleum*, *Leucaena leucocephala*, and *Lotus halophilus* (Berck et al., 1999). NodX functions as an acetyltransferase which acetylates the reducing end of Nod factors (Berck et al., 1999). While most of *R. leguminosarum* bv. *viciae* strains cannot nodulate *Pisum sativum* cv. Afghanistan, *R. leguminosarum* bv. *viciae* strain TOM that harbors NodX can overcome this and nodulate Afghanistan peas (Firmin et al., 1993). NodZ of *Bradyrhizobium* participates in fucosylation (Stacey, 1995, Lopez-Lara et al., 1996). Enzymes encoded by *nodPQ* of *Rhizobium tropici* CIAT899 are involved in the synthesis of adenosine 5′-phosphosulfate (APS) and 3′-phosphoadenosine 5′-phosphosulfate (PAPS), which NodH uses to transfer a sulfate group to the Nod factors (Folch-Mallol et al., 1998, Roche et al., 1991). After biosynthesis, Nod factors are secreted into the rhizosphere and perceived by the host receptor complex.

Receptor kinases located on the plasma membrane of root hairs recognize the Nod factors. Two Nod factor receptors have been identified in legumes. They are called *Lj*NFR1 and *Lj*NFR5 in *L. japonicus* (Madsen et al., 2003, Radutoiu et al., 2003), *Mt*NFP and *Mt*LYK3 in *Medicago truncatula* (Amor et al., 2003, Limpens et al., 2003, Arrighi et al., 2006), and *Ps*SYM10 and *Ps*SYM2A in *P. sativum* (Geurts et al., 1997, Madsen et al., 2003). These receptors are LysM domain-containing receptor kinases (LysM-RK), which have three extracellular lysin motifs (LysM1, LysM2, and LysM3), a transmembrane domain, and a kinase or pseudokinase domain (Wong et al., 2019). The extracellular LysM domain binds the chitin core of the Nod factors (Fliegmann & Bono, 2015, Murakami et al., 2018, Bozsoki et al., 2020), and the substitutions of Nod factors are discriminated by host receptors (D'Haeze & Holsters, 2002, Downie, 1998, Fliegmann & Bono, 2015, Roche et al., 1996). Transgenic introduction of the LysM-RKs from *L. japonicus* into *M. truncatula* and amino acid changes in the LysM domain of the receptor kinase result in the loss of Nod factor-induced signaling or the alteration of the rhizobial symbiosis partner (Amor et al., 2003, Madsen et al., 2003, Radutoiu et al., 2007). For

instance, introducing *Lj*NFR1 and *Lj*NFR5 into *M. truncatula* enables *M. truncatula* to nodulate with *Mesorhizobium loti,* the symbiont of *L. japonicus* (Radutoiu et al., 2007). Replacement of regions II and IV of *Lj*NFR1 with the respective *Lj*CERK6 regions leads to a dramatic decrease in *M. loti* Nod factor binding (Bozsoki et al., 2020). Moreover, the LysM2 of NFR5 shows the highest diversity in LysM domains among NFR1 and NFR5 homologs among investigated plant species (Madsen et al., 2003). A conformational change in LysM2 upon *M. loti* Nod factor binding has been observed (Sorensen et al., 2014). A change in Leu118 of LysM2 alters the Nod factor binding affinity of *M. loti* to *R. leguminosarum* bv. *viciae* strain DZL at the entry or exit of the groove of NFR5 (Radutoiu et al., 2007). This evidence supports the decisive role of NFR1 and NFR5 in symbiotic partner selection based on Nod factor perception.

3.2 Nod factor perception induced downstream signaling pathway

After Nod factor perception, calcium accumulates on the root hair, which prompts the calcium influx in the cytoplasm and calcium oscillations, also called calcium spiking, in the nucleus of epidermal root hair cells (Fliegmann & Bono, 2015). In *Lotus*, these calcium oscillations are activated after Nod factor perception by a heterocomplex formed by *Lj*NFR1, *Lj*NFR5, and a leucine-rich repeat (LRR) receptor kinase, the Symbiosis Receptor Kinase (*Lj*SYMRK) (Antolin-Llovera et al., 2014). The kinase domain of *Lj*NFR1 interacts with an E3 ubiquitin ligase and RGS (regulator of G-protein signaling) proteins (Choudhury & Pandey, 2015). The G-protein signaling pathway might function in stimulating the production of a secondary messenger, which is unidentified yet. Secondary messenger further leads to Ca^{2+} release in the nucleus (Charpentier, 2018). An additional epidermal LysM receptor controlling calcium spiking was identified in 2018, which is named NFRe. NFRe keeps the normal interval of calcium spiking in the root hairs. It phosphorylates NFR5, in return regulating its downstream signaling. Mutations in NFRe result in extended calcium spiking interval, lessened transcriptional response, and decreased nodule number upon rhizobia infection (Murakami et al., 2018).

The generation of calcium spiking involves three nucleoporin components – *Lj*NUP85, *Lj*NUP133, and *Lj*NENA (Kanamori et al., 2006, Saito et al., 2007, Groth et al., 2010), ion channels located on the nuclear membrane – *Lj*CASTOR, *Lj*POLLUX (Imaizumi-Anraku et al., 2005, Charpentier et al., 2008) and cyclic nucleotide-gated channels (Charpentier et al., 2016). The signal produced by calcium spiking is proposed to be decoded by the nuclear calcium-calmodulin kinase CCaMK, which interacts with and phosphorylates a coiled-coil protein, CYCLOPS (Yano et al., 2008, Singh & Parniske, 2012). The phosphorylated CYCLOPS acts as a DNA-binding transcriptional activator, transactivating *Nodule Inception* (*NIN*) that targets

hundreds of genes for rhizobial infection and nodule formation (Yano et al., 2008, Singh et al., 2014, Liu et al., 2019). Mutations in *NIN* of *L. japonicus*, *M. truncatula*, and *P. sativum* abort nodule formation (Schauser et al., 1999, Borisov et al., 2003, Marsh et al., 2007). In addition, *Ethylene Response Factor Required for Nodulation1* (*ERN1*), Nuclear Factor-Y (NF-Y) subunit gene *NF-YA1* (Laporte et al., 2014, Cerri et al., 2017), and two GRAS-domain-containing proteins, Nodulation Signaling Pathway1 (NSP1) and NSP2, are responsible for transcriptional reprogramming downstream of CCaMK (Kalo et al., 2005, Smit et al., 2005, Heckmann et al., 2006). It has been suggested that NF-YAs act downstream of *NIN* and *CCaMK*, mediating the NSP1/2 complexes to induce ERN1 (Laloum et al., 2014). Collectively, the perception of Nod factors elicits downstream regulatory pathways operating gene expression for rhizobial infection at root epidermis and distal cell divisions for nodule organogenesis in the root cortex (Soyano & Hayashi, 2014, Roy et al., 2020).

3.3 Exopolysaccharides and their perception by the hosts

In addition to Nod factors, rhizobial surface polysaccharides are relevant in symbiosis establishment as well. The surface polysaccharides serve as the backbone of biofilm, protective capsules, and signals for host recognition (Janczarek et al., 2015). Cyclic glucans, capsular polysaccharides, K-antigen capsular polysaccharides, lipopolysaccharides, and exopolysaccharides (EPS) are the five main rhizobial surface polysaccharides. Among them, the symbiotic function of EPS has been better studied (Acosta-Jurado et al., 2021).

EPS are poorly connected to the cell surface and secreted into the environment in a significant proportion (Acosta-Jurado et al., 2021). They are species- or strain-specific and consist of repeating subunits of common monosaccharides decorated with additional non-carbohydrate residues, such as acetyl, pyruvyl, succinyl, and methyl groups (Acosta-Jurado et al., 2021, Janczarek et al., 2015). For instance, rhizobia strains that are close in phylogeny share a similar basis of EPS but their EPS have different decorations. While the EPS of *R. leguminosarum* bv. *viciae* are composed of octasaccharides with D-glucuronic acid residues, the EPS of *R. leguminosarum* bv. *trifolii* strains is generally composed of octasaccharides with D-glucose, D-glucuronic acid, and D-galactose residues and modified by acetyl and pyruvyl groups (Acosta-Jurado et al., 2021, Marczak et al., 2017).

Perception of rhizobial EPS molecules is required for effective infection and is conducted by the LysM receptor Exopolysaccharide Receptor 3 (EPR3) in *L. japonicus* (Kawaharada et al., 2015). *M. loti* R7A with a mutation in *exoU*, secreting penta-glycan instead of O-acetylated acidic EPS, fails to infect its host *L. japonicus* Gifu, whereas increased infection by

M. loti R7A*exoU* is observed in *epr3* mutants (Kawaharada et al., 2015). Similarly, *exoY210* and *exoH225* mutants in *Sinorhizobium meliloti* 1021 fail to infect its host *Medicago sativa* due to the deficient production in succinoglycan (EPS II) and the production of symbiotically dysfunctional succinoglycan without succinyl modification, respectively (Cheng & Walker, 1998). Defects in EPS also affect the nitrogen fixation of nodules. Mutations in the EPS biosynthesis gene *pss1* of *R. leguminosarum* bv. *trifolii* ANU437 and *exo1* of *R. leguminosarum* bv. *viciae* ANU54 lead to the formation of non-nitrogen-fixing nodules on clover roots (Rolfe et al., 1996). Overall, EPS recognition is one of the factors, which determines symbiotic compatibility between host plants and rhizobia.

3.4 Effectors and their perception by the hosts

Several rhizobia strains secrete effectors into the host cytoplasm via Type III secretion system (T3SS), which modulates the symbiosis signaling pathways (Miwa & Okazaki, 2017). The T3SS of rhizobia strains, including *E. fredii* NGR234 (Freiberg et al., 1997), *E. fredii* HH103 (de Lyra Mdo et al., 2006), *B. elkanii* USDA61 (Okazaki et al., 2009), and *M. loti* MAFF303099 (Kaneko et al., 2000), have been identified. The effectors delivered by T3SS have been demonstrated in pathogenic bacteria, which are used to suppress host immune response (Macho & Zipfel, 2015).

Rhizobial T3SS effectors take part in symbiotic compatibility determination (Miwa & Okazaki, 2017). Two examples of T3SS contributing to symbiotic compatibility are found between *Lotus* accessions and broad-host-range rhizobia strains, *B. elkanii* USDA61 and *E. fredii* HH103. The T3SS of *B. elkanii* USDA61 induces infection inhibition in *L. japonicus* Gifu, while the infection is allowed in *L. burttii* and *L. japonicus* MG-20. This incompatibility on rhizobial infection in *L. japonicus* Gifu is due to the presence of NopF protein, a T3SS effector of *B. elkanii* USDA61 (*Be*NopF) (Kusakabe et al., 2020). Introducing *BeNopF* into the *Lotus*-compatible strain *M. loti* MAFF303099 results in infection incompatibility in *L. japonicus* Gifu (Kusakabe et al., 2020). *E. fredii* HH103 induces ineffective nodules on *L. japonicus* Gifu, whereas the induced nodules on *L. burttii* are nitrogen-fixing (Sandal et al., 2012). Mutants in T3SS, which cannot secrete the Nop effectors are able to induce nitrogen-fixing nodules on *L. japonicus* Gifu and *L. burttii*. Among the eight identified Nop effectors, NopC is the key determinant. All *L. japonicus* Gifu plants inoculated with *E. fredii* HH103 *nopC* mutant formed nitrogen-fixing nodules and presented a significant growth-promotion effect, whereas *L. japonicus* Gifu plants inoculated with other T3SS mutants showed less nitrogen-fixing nodules and shoot fresh weight (Jimenez-Guerrero et al., 2020). These results demonstrate that the T3SS determines the symbiotic compatibility in a host genotype-dependent manner.

Rhizobial T3SS is also involved in nodulation processes using a Nod factor-independent mechanism (Okazaki et al., 2013). While *B. elkanii* USDA61 and its Nod factor-deficient mutant can still nodulate *Glycine max* En1282 (an *nfr1*-mutant soybean variety), the T3SS-deficient *B. elkanii* USDA61 fails to nodulate this variant (Okazaki et al., 2013). The expression of *early nodulin 40* (*ENOD40*) and *NIN*, two nodulation-specific genes, is increased in the roots of En1282 inoculated with *B. elkanii* USDA61 but not with its T3SS mutant (Okazaki et al., 2013). These findings suggest that T3SS bypasses the Nod factor-dependent nodulation.

3.5 Plant selectivity over rhizobia strains

Legumes respond to different rhizobia species dissimilarly. Some plants nodulate with a broad spectrum of rhizobia, while others have a limited range. The difference within plant species after a specific rhizobia infection represents the symbiotic compatibility between host and rhizobia (Firmin et al., 1993). For example, *Phaseolus vulgaris* has been considered a non-selective host (Michiels et al., 1998). It nodulates with many species in the genera *Rhizobium*, *Ensifer* and *Pararhizobium*, and a minority in the genus *Bradyrhizobium* (Shamseldin & Velazquez, 2020). *Sophora flavescens* is a promiscuous host that can be nodulated by *Rhizobium yanglingense*, *Mesorhizobium amorphae* and *E. fredii*, but not *Bradyrhizobium diazoefficiens*. In contrast, only *E. fredii* and *B. diazoefficiens* can nodulate *G. max* (Liu et al., 2018b). The compatibility between *Lotus* and rhizobia is relatively restrictive. While *Lotus uliginosus, Lotus subbiflorus*, and *Lotus angustissimus* are effectively nodulated by *Bradyrhizobium* (Irisarri et al., 1996, Cooper et al., 1985), *Lotus corniculatus, Lotus tenuis* and *L. japonicus* are effectively nodulated by *Mesorhizobium* species (Saeki & Kouchi, 2000, Jarvis et al., 1982). Even though strains in *Rhizobium, Ensifer*, and *Aminobacter* are reported to nodulate *Lotus* spp. different *Lotus* accessions possess different compatibility with each rhizobial genus (Lorite et al., 2018). These findings show the variety of symbiotic compatibility is broadly spread in legumes species and diverse between host species and rhizobia strains.

The perception of cognate molecules determines the symbiotic compatibility between given pairs of rhizobia and plants. One of the earliest steps is the perception of rhizobial Nod factors mentioned in the previous section, which initiates nodule organogenesis (Limpens et al., 2003, Oldroyd & Downie, 2004, Oldroyd et al., 2011, Wang et al., 2012). Rhizobia strains with mutations in *nodC*, *nodM*, and *nodE*, which affect the Nod factor structure, fail to nodulate the host plant (Liu et al., 2018b). In addition to Nod factors, rhizobial EPS has been proposed to play a role in rhizobia-legume compatibility. A two-step receptor-mediated recognition of compatible rhizobia has been proposed, in which Nod factor perception is followed by the EPR3-EPS perception to discriminate between compatible and incompatible rhizobia

(Kawaharada et al., 2015). The effectors secreted by T3SS also contribute to host specificity. *G. max* plants expressing Rfg1, an NBS-LRR disease resistance protein, block the nodulation with *E. fredii* USDA257 wild type but not the T3SS knockout mutant (Yang et al., 2010). These studies demonstrate the perception of Nod factors, rhizobial EPS, and effectors is involved in determining symbiotic compatibility, however, the underlying mechanisms remain obscure.

3.6 Rhizobia infection and nodule organogenesis

Nodules that arise from cortical cell divisions are initiated downstream Nod factor signaling in parallel with the rhizobial infection process (Oldroyd et al., 2011). After Nod factor perception, NIN targets the promoter of *NF-YA1* and *NF-YB1* to promote cortical cell divisions (Soyano et al., 2013). NIN also targets *Asymmetric Leaves 2-Like/Lateral Organ Boundaries domain 16* (*ASL18*/*LBD16*) (Soyano et al., 2019), which controls lateral root generation in response to auxin (Okushima et al., 2007). Similar spatial expression patterns are shown between *NIN*, *NF-Y* subunits, and *ASL18*/*LBD16* in nodule primordia (Soyano et al., 2013, Soyano et al., 2019). A high degree of overlap in gene expression changes and developmental programs is also observed between lateral roots and nodules (Schiessl et al., 2019). Furthermore, co-expression of *NF-YA1*, *NF-YB1*, and *ASL18/LBD16* increases lateral root density up to six-fold (Soyano et al., 2019). These results suggest a shared regulatory pathway between nodule organogenesis and lateral root development downstream of NIN.

Phytohormones, especially cytokinin and auxin, are involved in the early stage of nodule development. A cytokinin receptor encoded by *Cytokinin Response 1* (*CRE1*) in *Medicago* and *Lotus histidine kinase* (*LHK1)* in *Lotus*, positively contributes to cortical cell division and nodule formation (Gonzalez-Rizzo et al., 2006, Murray et al., 2007, Tirichine et al., 2007b, Vernie et al., 2015). Repressing *CRE1* expression by RNA interference (RNAi) makes the root insensitive to cytokinin, and causes an increase in lateral roots and a dramatical decrease in nodulation (Gonzalez-Rizzo et al., 2006). In contrast, a gain-of-function and cytokinin-hypersensitive mutant version of *LHK1*, named *snf2-2*, generates spontaneous nodules even without rhizobia (Tirichine et al., 2007b). Hence, the active cytokinin signaling pathway is not only essential but also sufficient for nodule organogenesis.

The cytokinin signaling is responsible for the CRE1-dependent alteration of polar auxin transport, which suppresses the expression of PIN auxin efflux transporter (Oldroyd et al., 2011, Plet et al., 2011, Ariel et al., 2012). This results in a low auxin concentration in the cortex, which promotes nodule primordia formation (Grunewald et al., 2009, Oldroyd et al., 2011). In the *cre1* mutant, the auxin transport fails to respond to rhizobia inoculation and root nodulation is

reduced (Ng et al., 2015). In addition, cytokinin can induce NIN expression, which activates *ASL18*/*LBD16* that then stimulates auxin biosynthesis through inducing *STYLISH* (*STY*) and *YUCCAs* (*YUC*) expression for nodule development (Schiessl et al., 2019). Overall, cytokinin signaling coordinates with polar auxin transport to regulate nodule organogenesis.

4 Regulation of nodulation

To accommodate rhizobia for nitrogen fixation, the host plants need not only to generate nodules on the roots but also to provide photosynthesis products to sustain the living of rhizobia in the nodules. These processes consume energy and resources, which are costly for the hosts. Therefore, plants develop mechanisms to control nodule formation and rhizobia infection (Tanabata & Ohyam, 2014, Wang et al., 2018a, Ferguson et al., 2019). To regulate nodulation, plants coordinate endogenous and environmental cues (Capoen et al., 2010, Liu et al., 2018a, Nishida et al., 2020, He et al., 2021).

4.1 Phytohormone regulation

Phytohormones play crucial roles in every stage during plant growth (Foo et al., 2019). Studies using mutants and exogenous application of hormones indicate that nodulation is under the control of phytohormone-inclusive networks (Lorteau et al., 2001, Lin et al., 2020). Phytohormones fine-tune the nodulation process via hormone cross-talks or in response to environmental factors, although the detailed mechanisms are unclear (Lin et al., 2020). The balance of phytohormones seems to be critical for nodulation, as they impact rhizobia infection, nodule organogenesis, and nodule senescence (Foo et al., 2019, Dolgikh et al., 2019).

4.1.1 Gibberellin and nodule formation

Gibberellins (GAs) comprise a group of diterpenoid carboxylic acids that function as growth regulators, facilitating organ expansion and modulating developmental processes (Hedden & Thomas, 2012). Contradictory effects of GA application on nodulation have been reported in *L. japonicus.* In one study, GA promotes nodule formation in a range of 10^{-7} M to 10^{-4} M (Kawaguchi et al., 1996). In another study, GA application inhibits both IT and nodule formation in a range of 10^{-8} M to 10^{-6} M. Adding the biosynthetic inhibitor of GA, uniconazole P, promotes both IT and nodule formation at 10^{-7} M (Maekawa et al., 2009). In *Medicago*, GA has a negative effect on nodulation (Fonouni-Farde et al., 2016b, Jin et al., 2016). Yet, a study in GA-deficient mutants of pea, *na-1*, *ls-1*, and *lh-2*, shows a reduction in nodule organogenesis. Especially in *na-1*, which produces less GA_1, forms the lowest number of nodules among the three mutants.

This reduction in nodule number can be restored by GA application or grafting wild-type shoots or roots to the mutant plant. In addition, ethylene production is doubled in *na-1* mutants. This finding indicates an influence of GA concentration on ethylene production and ethylene may also contribute to the reduced nodule number (Ferguson et al., 2011). However, high GA concentration inhibits nodule formation in wild-type plants (Ferguson et al., 2005). The responses to GA seem to be species-dependent. The inconsistent results to GA addition can be explained by variations in growth conditions, artifacts, or applied GA type and concentration (Hayashi et al., 2014, Lin et al., 2020).

SLY1 and DELLA proteins are involved in the GA signaling pathway and regulate nodulation (Figure 3A) (Maekawa et al., 2009, Fonouni-Farde et al., 2016b, Jin et al., 2016). In *L. japonicus*, overexpressing *LjSly1a* and a gain-of-function mutant *sly1*-d present a reduction in nodule number, suggesting that constitutive GA signaling represses nodule organogenesis (Maekawa et al., 2009). In pea and *Medicago*, DELLA proteins have a positive role in nodule organogenesis. Loss of DELLA proteins causes a decrease in nodule number (Ferguson et al., 2011, Fonouni-Farde et al., 2016b, Jin et al., 2016). Studies in *Medicago* indicate that *Mt*DELLA proteins can interact with *Mt*NSP2, *Mt*NF-YA1, and *Mt*IPD3 (an orthologue of *Lotus* CYCLOPS). The interaction between DELLA and NSP2, NF-YA1, and IPD3 promotes the expression of nodulation genes *MtERN1* and *MtNIN* for nodule organogenesis (Fonouni-Farde et al., 2016a, Fonouni-Farde et al., 2016b, Jin et al., 2016).

4.1.2 Cytokinin and autoregulation of nodulation

Cytokinin was discovered by its regulatory role in cell division (Arora et al., 1959, Hirsch et al., 1997, Fang & Hirsch, 1998), which is essential for nodule organogenesis (Dolgikh et al., 2019). Cytokinin addition results in the formation of nodule-like structures on *L. japonicus* Gifu (Heckmann et al., 2011). While a gain-of-function mutant in the *snf2* allele of the *LjLhk1* cytokinin receptor gene forms spontaneous nodules in the absence of rhizobia, a loss-of-function mutant *lhk1-1* fails to form nodules in the presence of rhizobia (Tirichine et al., 2007a, Heckmann et al., 2011). Application of cytokinin also leads to increased nodulation in *Sesbania rostrata* and *M. sativa* (Lin et al., 2020). At concentrations lower than 10^{-6} M, cytokinin promotes nodule formation in *P. sativum* cv. Sparkle, whereas higher concentrations reduce nodulation (Lorteau et al., 2001). Similarly, cytokinin promotes nodulation in *L. japonicus* MG-20 at 10^{-8} M, while concentrations higher than 10^{-7} M reduce nodule numbers (Figure 3A) (Sasaki et al., 2014). Therefore, cytokinin regulation is important as it can affect nodulation both positively and negatively.

In addition to the counteracting role in nodule development with auxin (section 3.6), cytokinin is also involved in the so-called autoregulation of nodulation (AON) (Figure 3B). AON is a negative feedback loop, acting long-distance between root and shoot to suppress the nodule organogenesis on the root (Heckmann et al., 2011, Nishida & Suzaki, 2018). As a result of Nod factor perception, peptides belonging to the Clavata3/Embryo Surrounding Region-Related (CLE) Family are produced from the roots. These peptides include CLE-Root Signal 1 (CLE-RS1), CLE-RS2, and CLE-RS3 in *L. japonicus*, CLE12/13 in *Medicago*, and Rhizobia Induced CLE1 (RIC1) and RIC2 in *G. max* (Nishida & Suzaki, 2018, Roy et al., 2020). CLE peptides consist of 12-13 amino acids, which are cleaved from a prepropeptide of about 100 amino acids in length (Hastwell et al., 2015). These CLE peptides are transported via the xylem and perceived in the shoot by a receptor complex, inducing HAR1 in *L. japonicus* and its orthologues NARK in *G. max*, and SUNN in *M. truncatula* (Ferguson et al., 2019). In *L. japonicus*, this CLE-RS1/2-HAR1 complex initiates cytokinin biosynthesis by activating an isopentenyl transferase (IPT)-encoding gene *IPT3* and leads to downregulation of miR2111. The shoot-derived cytokinin is then translocated to the root and block nodule development (Magori et al., 2009, Takahara et al., 2013, Sasaki et al., 2014). Downregulation of miR2111 results in higher expression of Too Much Love (TML), a F-box protein in the root, which also inhibits nodule formation (Lin et al., 2020). Similarly, exogenous application of cytokinin on the wildtype *L. japonicus* MG-20 shoot and grafting *IPT3*-overexpressed shoot onto roots of wildtype *L. japonicus* MG-20 both result in nodulation inhibition on roots (Sasaki et al., 2014). This inhibitory effect is eliminated in the *har1-7* mutant, which shows a hypernodulation phenotype. The *tml-1* mutant can escape this inhibitory effect as well (Takahara et al., 2013, Sasaki et al., 2014). Furthermore, rhizobia-induced cytokinin can regulate the CLE-HAR1 negative feedback pathway from roots in a NIN-dependent manner (Laffont et al., 2020, Lin et al., 2020). Altogether, AON regulates nodule formation by CLE-HAR1-TML regulatory module via long-distance signaling.

4.1.3 Ethylene regulation of nodule formation

Ethylene is a gaseous phytohormone, which is increased within 6 h post rhizobia inoculation (Reid et al., 2018). It functions negatively on nodule numbers and IT formation in various legumes (Oldroyd et al., 2001). Ethylene Insensitive 2 (Ein2) has been identified as a central regulator of ethylene signaling (Ju & Chang, 2015). *M. truncatula* and *L. japonicus* e*in2* mutants become ethylene-insensitive. The *Mtein2* (formerly called *sickle*) presents both hypernodulation and hyperinfection phenotypes (Penmetsa & Cook, 1997, Penmetsa et al., 2008). In *L. japonicus*, *LjEin2* is duplicated and functions redundantly. Mutations in both *ein2* homologs lead to hypernodulation, whereas either *ein2-1* or *ein2-2* single mutant is only insensitive to ethylene without hypernodulation (Miyata et al., 2013, Reid et al., 2018).

Blocking ethylene perception or biosynthesis respectively by silver ion or aminoethoxyvinyl glycine (AVG) increases nodulation (Peters & Crist-Estes, 1989, Nukui et al., 2000, Oldroyd et al., 2001, Heckmann et al., 2011). On the contrary, adding ethylene precursor, 1-aminocyclopropane-1-carboxylic acid (ACC), reduces nodulation (Nukui et al., 2000, Oldroyd et al., 2001). Applying silver thiosulfate (STS, Ag^+ doner) and AVG increases nodule number in *M. sativa*, *Macroptilium atropurpureum*, and *L. japonicus* upon rhizobia inoculation. In contrast, reduced IT and nodulation are observed when 1 μM ACC is applied (Nukui et al., 2000). In *M. truncatula*, the number of infections and nodules raises progressively when decreasing the concentration of ACC from 10, 1 to 0.1 μM and increasing the concentration of AVG from 0.01, 0.1 to 1 μM (Oldroyd et al., 2001). Inhibitory effects of ethylene are also found in *M. sativa* (Peters & Crist-Estes, 1989), *P. sativum* cv. Sparkle (Lee & Larue, 1992) and *Trifolium repens* (Goodlass & Smith, 1979). These findings imply a negative role of ethylene in nodule formation (Figure 3A).

The crosstalk between ethylene and cytokinin inhibits nodule formation. The suppression of both nodule formation and Ca^{2+} spiking in *M. truncatula* root hairs is observed in response to Nod Factors after ACC application (Oldroyd et al., 2001). Notably, ACC application represses the spontaneous nodulation in *snf1* (a gain-of-function CCaMK point mutation line), *snf2-1* and *snf2-2*, indicating that the ethylene has an inhibitory effect on cytokinin signaling (Tirichine et al., 2006). Moreover, a feedback loop is proposed between ethylene and cytokinin responses. A positive correlation between ethylene production and exogenous cytokinin application was revealed (Lorteau et al., 2001). The induction of ACC synthases upon Nod factor application is dependent on *CRE1* in *M. truncatula* (van Zeijl et al., 2015), and increased ethylene emission is observed in uninoculated *L. japonicus snf2* (Reid et al., 2018). This indicates a positive regulatory role of cytokinin in ethylene biosynthesis and signaling. However, an ethylene-independent regulation on nodulation has been reported in soybean (Schmidt et al., 1999, Nukui et al., 2000). Overall, the ethylene-dependent inhibition coordinates with cytokinin to control nodule formation.

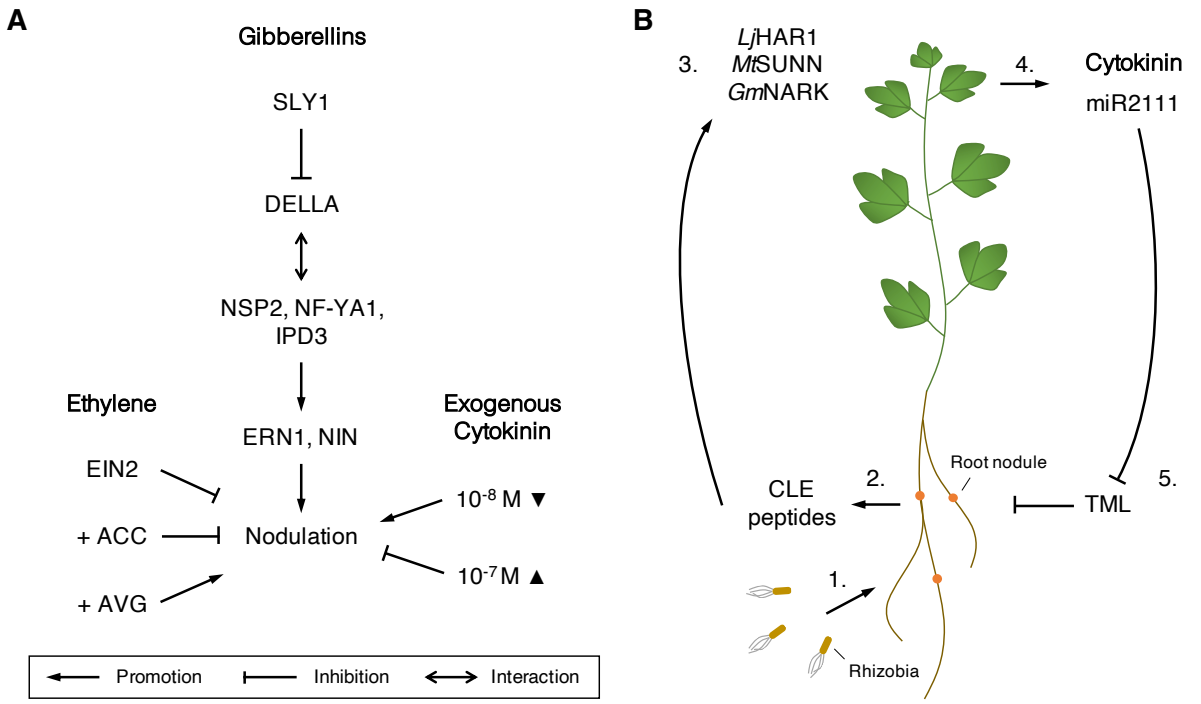


Figure 3. Phytohormones in the regulation of nodulation. (A) Gibberellin, ethylene, and cytokinin regulation on nodulation. SLY in the gibberellin signaling inhibits DELLA, which can interact with NSP2, NF-YA1, and IPD3. This interaction promotes the expression of nodulation genes *ERN1* and *NIN*, which mediated nodule formation signaling pathway. Ein2 is the central regulator of ethylene signaling, which represses nodule formation. Application of the ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) inhibits nodulation, while the ethylene biosynthesis inhibitor aminoethoxyvinyl glycine (AVG) promotes nodulation. The effect of exogenous cytokinin application is presented at the bottom-right of the figure, which is concentration- and plant species-dependent (As example is shown the effect of cytokinin addition on *L. japonicus* MG-20). (B) Autoregulation of nodulation. After root perceiving rhizobia Nod factors (1), CLE peptides are produced (2). The CLE peptides are transported to shoots and recognized by receptors there (3). The CLE-receptor complex initiates cytokinin production and leads to downregulation of miR2111 in the shoots (4) that inhibits Too Much Love (TML) in the roots.

4.2 Drought and waterlogging regulate nodulation

Plants face various fluctuations of environmental alteration or stresses during their lifetimes. Heat, cold, and drought conditions are common abiotic stresses (Fedoroff et al., 2010). Among them, variation in soil moisture can lead to water stresses accompanied by changes in soil salinity and nutrient deficiency (Saijo & Loo, 2020). These abiotic stresses influence the molecular activity and physiology of both plants and rhizobia (Bordeleau & Prévost, 1994). Plants respond to the stresses by altering the production of phytohormones, antioxidants, reactive oxygen species (ROS) and adjusting the growth of roots and shoots (Bao et al., 2014, Verma et al., 2016). Rhizobia also alter the production of their EPS and osmoprotective molecules in response to stresses, which affect rhizobial colonization and survival (Janczarek,

2011, Vriezen et al., 2007). Hence, the symbiotic compatibility is changed under stress conditions.

4.2.1 The effects of soil moisture on the host, rhizobia, and their symbiosis

Water content in growth substrate is critical for plants and rhizosphere microbes (Borowik & Wyszkowska, 2016, Scharwies & Dinneny, 2019). For rhizobia, limited water availability can lead to the accumulation of salts, increase osmotic stress, and impair cell metabolism (Vriezen et al., 2007). A decrease in substrate utilization (Griffiths et al., 2003) and an exponential decline in nitrification rates of nitrifying bacteria are observed under low water availability (Stark & Firestone, 1995). The water content also has implications on the bacteria population (Griffiths et al., 2003, Borowik & Wyszkowska, 2016) and motility (Soby & Bergman, 1983). When the water content varies in the growth substrate, the number of rhizobia colonies (Borowik & Wyszkowska, 2016) and the spreading rate of *E. meliloti* (Soby & Bergman, 1983) fluctuates. These alterations may be detrimental for rhizobia to establish symbiosis with hosts.

Legumes are sensitive to water deficiency and waterlogging, although they can tolerate a short-term water excess and shortage (Sprent, 1984, Bordeleau & Prévost, 1994). In *L. japonicus*, drought represses the expression of genes involved in photosynthetic metabolism and induces the expression of genes involved in the antioxidant response, indicating increased oxidative stress (Betti et al., 2012). Plants also adapt to water stress conditions by altering the growth of the root system and root hairs (Bao et al., 2014). The abnormal growth of root hairs restricts the root hair infection under drought conditions (Lie, 1981). Waterlogging often leads to hypoxia or even anoxia in roots due to the quick O_2 consumption and the slow O_2 diffusion, reducing the root growth and nutrient uptake (Striker & Colmer, 2017). The semiaquatic plant *S. rostrata* and its rhizobia symbiont *Azorhizobium caulinodans* have evolved a dual infection mechanism to bypass this limitation. Under non-flooding conditions, rhizobia enter via IT formed in root hair. Instead, crack entry is used upon flooding since root hair growth is inhibited (Goormachtig et al., 2004).

Water stress causes reductions in nitrogen fixation and respiration in nodules (Coba de la Peña & Pueyo, 2011). The salinity stress is often coupled with water deficiency and detrimental for nodules (Bordeleau & Prévost, 1994, Bruning & Rozema, 2013). A decrease in nitrogenase activity is found in nodules suffering from drought and salinity (Naya et al., 2007, Lopez et al., 2008), which is a consequence of defects in oxygen permeability around nodules (Del Castillo et al., 1994, Naya et al., 2007). Pea (Minchin & Pate, 1975), alfalfa (Arrese-Igor et al., 1993), and soybean (Sung, 1993) grown under hypoxic conditions show a reduction in nodule weight

and the number. A study shows that the severity of hypoxia determines its effect on nodulation. For *Medicago* under moderate hypoxia conditions with 4.5% O_2, nodulation is not affected. Yet, the nodule number decreases when a high hypoxia condition (0.1% O_2) is applied (El Msehli et al., 2016). Moreover, nitric oxide (NO) produced under hypoxia conditions represses the nitrogen fixation by inhibiting the expression and activity of the nitrogenase (Salas et al., 2019).

Even some legume species are flood-tolerant, they still possess few morphological and physiological adaptations during nodulation (Striker & Colmer, 2017, Pucciariello et al., 2019). Two strategies have been adopted by *Lotus* to bypass submergence. While *L. japonicus* plants allocate carbon towards shoot elongation to escape from below the water, *L. corniculatus* and *L. tenuis* are quiescent in submergence and show higher recovery growth in de-submergence (Striker et al., 2012). An exception was discovered in *Smilax herbacea* (carrion flower), which benefits from waterlogging. Carrion flower presents taller plants, higher biomass, and more nodules under flooding conditions than plants grown on dry land. Biochemical adaptation is presumable to be the cause, including the presence of polyhydroxybutyrate crystals in bacteroids, reduction in asparagine, and increased aspartate aminotransferase for assimilation of reduced nitrogen (Krishnan et al., 2019). Overall, water stress is often detrimental for nodulation and tolerance is species-dependent.

4.2.2 Water stress impacts the nodulation via the action of abscisic acid

In drought, abscisic acid (ABA) biosynthesis is activated and its degradation is restrained, resulting in ABA accumulation (Verma et al., 2016). ABA biosynthesis genes, inclusive of genes encoding the zeaxanthin epoxidase (ZEP), aldehyde oxidase (AAO), 9-cis-epoxycarotenoid dioxygenase (NCED), and molybdenum cofactor sulfurase, are upregulated (Zhu, 2002). ABA contributes to nodulation negatively. Exogenous ABA application causes a decrease in nodule number and weight in *G. max* (Cho & Harper, 1993, Bano & Harper, 2002). Similar results are obtained when ABA is applied to *Vigna radiata* (Farooq & Bano, 2007), *M. truncatula* (Ding et al., 2008), *T. repens*, and *L. japonicus* (Suzuki et al., 2004). An ABA-hypersensitive mutant, *Mtsta-1*, exhibits an ABA-inhibitory effect on nodule formation as well (Ding et al., 2008). When NCED activity of *L. japonicus* MG-20 is inhibited by abamine, the ABA content is reduced and the nitrogen fixation is elevated. The *enhanced nitrogen fixation 1* (*enf1*) mutant, which comprises low endogenous ABA content, presents a 1.7 fold nodule number and 1.8 fold nitrogen fixation compared to the *L. japonicus* MG-20 wild type at 28 days post-inoculation (Tominaga et al., 2010). ABA also blocks rhizobia infection. For instance, white clover plants inoculated with *R. leguminosarum* bv. *trifolii* strain 4S and treated with ABA have a decreased number of curled root hairs, although the proportion of swelled root hairs is increased (Suzuki et al., 2004). In *Medicago*, the microcolonies of *S. meliloti* in the infection

foci and the number of ITs decrease with an increasing concentration of ABA (Ding et al., 2008). Conclusively, ABA has a detrimental effect on nodulation.

ABA has been shown to affect the early stage of root nodule symbiosis. Application of 1 mM ABA can suppress the calcium spiking induced by 1 nM Nod factor, which can be restored by ABA washout. ABA concentrations lower than 200 μM can extend the interval of calcium spiking, therefore reducing the spiking frequency in a given period. ABA also represses the expression of *peroxidase precursor 1*, *ENOD11*, and *ENOD40*, which are induced by rhizobia, Nod factor, and cytokinin, respectively (Ding et al., 2008). In *M. truncatula* with a transformed *abscisic acid insensitive 1* allele of *Arabidopsis*, plants are ABA-insensitive, possess a hypernodulation phenotype, and exhibit higher *ENOD40* induction after cytokinin treatment. Furthermore, ABA inhibits the spontaneous nodule organogenesis in *snf2* of *L. japonicus* (Ding et al., 2008). These results demonstrate that ABA regulates nodulation at multiple steps.

4.2.3 Reactive oxygen species impact the symbiosis

Both drought and waterlogging can lead to enhanced ROS production (Hasanuzzaman et al., 2020) and impact the symbiosis via the action of ROS. ROS play a significant role in controlling the legume-rhizobia symbiosis, especially at early stages (Damiani et al., 2016). A systemic ROS induction has been observed in soybean leaves at 30 minutes after *B. japonicum* inoculation (Fernandez-Gobel et al., 2019). A transient ROS burst has been discovered in root hairs a few seconds after Nod factors addition in *P. vulgaris*, which is inhibited by Nod factors after minutes (Cardenas et al., 2008). Inoculation of *S. meliloti* impaired in producing Nod factors eliminates the ROS production in *M. truncatula* (Ramu et al., 2002), which indicates that ROS generation is positioned downstream of Nod factor perception. The suppression effect of Nod factors on ROS generation and salicylic acid accumulation caused by defense response has been noticed in *M. truncatula* and *M. sativa* roots (Damiani et al., 2016). Hence, it is proposed that Nod factors trigger the primary ROS production for rhizobia infection (IT development and nodule formation) then inhibit the secondary ROS production to repress the defense response (Damiani et al., 2016).

The control of ROS scavenging affects the establishment of the symbiosis. During infection, O_2^- and H_2O_2 are generated in ITs and infected cells, while the genes encoding NADPH oxidases have been detected during IT progression as well as in nodules (Damiani et al., 2016). Suppressing the expression of *respiratory burst oxidase homologues* (*RbohA* and *RbohB*) in *P. vulgaris* attenuates infection, nodule organogenesis, and nitrogen fixation (Montiel et al., 2012, Arthikala et al., 2017). When *PvRbohB* is transiently knock-down by RNAi, the IT progression is impeded and ROS production is reduced coupled with the reduction in numbers

of nodule primordia and nodules (Montiel et al., 2012). In *PvRbohA* RNAi lines, the IT progression is aborted in root hairs without impairing cortical cell division. Thus, rhizobia cannot be released from IT and the nodules fail to fix nitrogen (Arthikala et al., 2017). Similar to *P. vulgaris*, superoxide anions accumulate in the IT and nodule primordia of *M. sativa* roots (Santos et al., 2001) and a ROS burst is observed in root cortical cells of *M. truncatula* (Peleg-Grossman et al., 2012). *MtRbohA* expresses specifically in the nitrogen-fixing zone of functional nodules and decreasing *MtRbohA* expression by RNAi results in a reduction in nitrogen fixation in *M. truncatula* (Marino et al., 2011). Altogether, these findings suggest that ROS and ROS scavenging are connected to nitrogen fixation in nodules, which require a tight control for symbiosis establishment.

4.2.4 Potential effect of water stress on nodulation due to the alteration in root exudate composition

The composition and quantity of the root exudates vary depending on the plant species, age, biotic and abiotic conditions (Haichar et al., 2008, Lombardi et al., 2018, Cesari et al., 2019). One important abiotic factor is the fluctuation in soil water content (Calvo et al., 2017). For instance, compounds such as ethanol, lactic acid, and alanine are secreted by plants in order to prevent cell damage from aerobic respiration under low O_2 conditions resulted from flooding (Hartman & Tringe, 2019), which may play roles in rhizobacteria selection. During flooding, *Aquaspirillum* sp. become the predominant rhizobacteria in the rhizosphere of poplar trees due to their ability to catabolize ethanol species secreted by poplar trees (Graff & Conrad, 2005). In contrast, the accumulation of phenolic compounds such as flavonoids is essential to prevent the damage caused by drought (Sharma et al., 2019). In white clover, the level of quercetin glycosides is increased by 111% in drought (Ballizany et al., 2012). As flavonoids are compounds that initiate rhizobial Nod factor production during root nodule symbiosis (Zhang et al., 2009, Abdel-Lateif et al., 2012), the alteration of flavonoids in root exudates has the potential to affect the symbiosis. Although water stress significantly affects the composition and quantity of root exudates (Henry et al., 2007), how the changes in the root exudates influence the rhizobacteria remains poorly understood (Bhattacharyya et al., 2021).

5 Natural variation in nodulation compatibility

Variation in symbiotic compatibility can be found between different species of legumes and rhizobia, which can also be found between host ecotypes and rhizobia strains (Granada et al., 2014). *M. loti* and *Bradyrhizobium* sp. strains nodulate *Lotus* species in a relatively specific manner (Hernández et al., 2005). While *L. corniculatus* and *Lotus glaber* establish symbiosis with *M. loti*, *L. subbiflorus* and *L. uliginosus* establish symbiosis with *Bradyrhizobium* strains.

These two symbiotic groups are incompatible to each other (Hernández et al., 2005). This finding suggests that the intraspecies genomic diversity in plants has an impact on symbiotic compatibility. However, the current understanding of the mechanism determining symbiotic compatibility remains limited (Wang et al., 2018b, Gossmann et al., 2012).

R. leguminosarum nodulates legumes within *Pisum*, *Lens*, *Lathyrus*, *Vicia*, and *Trifolium* genera (Fred et al., 1932). However, a *R. leguminosarum* strain Norway (*Rl* Norway) was isolated from *L. corniculatus*, and nodulates different *Lotus* species and ecotypes with variation in symbiotic compatibility (Gossmann et al., 2012). While *Mesorhizobium norvegicum* 10.2.2, a compatible symbiont of *Lotus* that was isolated from the same nodule as *Rl* Norway, nodulates *Lotus* species and ecotypes equally with nitrogen-fixing nodules on the root, *Rl* Norway induces non-functional, white nodules or bumps on *L. burttii*, *L. japonicus* MG-20, and *L. japonicus* Nepal (Gossmann et al., 2012, Kabdullayeva et al., 2020). The most striking phenotypic difference is between *L. burttii* and *L. japonicus* Gifu as *Rl* Norway induces white nodules on *L. burttii* but fails to induce nodules on *L. japonicus* Gifu (Gossmann et al., 2012). Moreover, *Rl* Norway induces the β-glucuronidase (GUS) activity in the *L. japonicus* Gifu T90 reporter line. Since the reporter line responds to rhizobia and Nod factor, this observation indicates that *L. japonicus* Gifu responds to *Rl* Norway, although it cannot generate nodules (Gossmann et al., 2012, Webb et al., 2000). Thus, *Rl* Norway is considered a sub-compatible symbiont of *Lotus*, presenting an accession-dependent nodulation compatibility.

As this diverse symbiotic compatibility is unique to *Rl* Norway and not prevalent between compatible partners, a system using *L. burttii*, *L. japonicus* Gifu, and *Rl* Norway to study symbiotic compatibility was established. In a previous study, recombinant inbred lines between *L. burttii* and *L. japonicus* Gifu were generated (Sandal et al., 2012) and phenotyped for Quantitative trait locus (QTL) mapping (Zarrabian et al., 2021). A QTL on chromosome 1 of *L. japonicus* Gifu close to the *NFR1* gene was identified, which co-segregated with the nodulation phenotype. In this QTL, three LysM-RKs were identified. Trans-complementation of the LysM-RK orthologs from *L. burttii* into *L. japonicus* Gifu did not grant nodulation in *L. japonicus* Gifu (Zarrabian et al., 2021). Altogether, how genetic variation contributes to symbiotic compatibility and whether additional factors play roles in symbiotic compatibility determination remain to be clarified.

Aim of the thesis

Molecular interaction between legumes and rhizobia is essential for establishing root nodule symbiosis (Oldroyd et al., 2011). Nod factor perception by plant NFR1 and NFR5 receptor complex is one determinant of symbiotic compatibility (Radutoiu et al., 2003, Fliegmann & Bono, 2015), where rhizobial EPS and effectors also play roles (Janczarek et al., 2015, Miwa & Okazaki, 2017). Abiotic factors, especially the water content in growth substrate, can lead to stress conditions and change the physiological activities in hosts and rhizobia. These changes can influence symbiotic compatibility by repressing the efficiency of nitrogen fixation or nodule formation (Lie, 1981, Bordeleau & Prévost, 1994, Pucciariello et al., 2019).

Variation in symbiotic compatibility between host and rhizobia species (Striker et al., 2012) also occurs between host ecotypes and rhizobia strains (Granada et al., 2014). The variation in symbiotic compatibility between compatible symbiont and hosts is subtle, while the variation in symbiotic compatibility between sub-compatible rhizobia and hosts is obvious. While compatible partners reach efficient nitrogen-fixing symbiosis, plants colonized by sub-compatible rhizobia encounter growth retardation, developmental limitation, and reduced crop production (Terpolilli et al., 2008, Friesen, 2012, Israel et al., 1986). Thus, it is important to clarify how genetic variation impacts symbiotic compatibility. As *Rl* Norway is a sub-compatible symbiont of *Lotus* that nodulates *L. burttii* but not *L. japonicus* Gifu, a system using *L. burttii* and *L. japonicus* Gifu, and *Rl* Norway to study symbiotic compatibility was established (Gossmann et al., 2012, Liang et al., 2019). A QTL was identified, which is linked to the nodulation phenotype (Zarrabian et al., 2021). However, the genes that contribute to this symbiotic compatibility require investigation.

In this study, I aimed to 1) identify the plant genes that contribute to the symbiotic compatibility in *Lotus*, and 2) determine how symbiosis compatibility is modulated by substrate parameters, such as moisture. To achieve the aims, an integrated approach, including phenotypic observations, transcriptomic sequencing, and molecular biology methods, was used.

Materials and methods

1 Plant growth, inoculation, and phenotyping

1.1 Bacteria cultivation

Mesorhizobium loti MAFF303099-*Ds*Red (*Ml* MAFF) and *Rhizobium leguminosarum* Norway-GFP (*Rl* Norway) were cultured at 28°C and 180 rpm for 2 days in tryptone yeast (TY) broth (Beringer, 1974). *Ensifer fredii* HH103 (*Ef* HH103) and *Allorhizobium undicola* LMGT (*Au* LMGT) were cultured at 28°C and 180 rpm for 2 days in yeast mannitol (YM) broth (Allen & Allen, 1950). Media were supplemented with selective antibiotics as follows: *Ml* MAFF, 25 µg ml^{-1} gentamicin; *Rl* Norway, 2 µg ml^{-1} tetracycline and 500 µg ml^{-1} streptomycin.

Escherichia coli strain TOP10 were grown in lysogeny broth (LB) medium (Bertani, 1951) at 37°C for 12-16 h. Media were supplemented with appropriate antibiotics (25 µg ml^{-1} Gentamycin or 100 µg ml^{-1} Spectinomycin). *Agrobacterium rhizogenes* strain AR1193 and *Agrobacterium tumefaciens* strain Agl1 were grown in yeast extract broth (YEB) medium (Vervliet et al., 1975) at 28°C for 2 nights. Media were supplemented with 50 µg ml^{-1} of carbenicillin, rifampicin, and kanamycin. Every *Agrobacterium* culture was freshly prepared in 3 ml YEB liquid medium using single colonies, and incubated at 28°C and 180 rpm for 2 nights. The information of the strains used in this work is listed in Table S1.

1.2 Plant materials

Lotus burttii B-303, *Lotus corniculatus* Leo (Georg Andreae GmbH), and eight *Lotus japonicus* accessions (*Lotus japonicus* Gifu B-129, MG-20, MG-70, MG-86, MG-119, MG-123, MG-135, and MG-136) were used to characterize the variations in nodulation phenotype upon rhizobia inoculation. Additionally, mutant lines in the *L. japonicus* Gifu background (*snf1-1* and *snf2-2*) were phenotyped after *Rhizobium leguminosarum* Norway inoculation under low and high substrate moisture conditions. To validate the functions of selected candidate genes, a total of 20 insertion-mutants of *L. japonicus* Gifu generated by *de novo* activation of the LTR element *Lotus retrotransposon 1* (LORE1, Malolepszy et al., 2016) were ordered from Lotus Base (https://lotus.au.dk) for *RLP1*, *RLP2*, and *RLP3*. Selected lines were used in experiments after propagation and genotyping. Moreover, *L. japonicus* Gifu and *RLP3* mutant line L30103870 were further used in trans-complementation assays. To generate stable mutant lines in the gene of interest, the CRISPR-Cas12a gene-editing method was performed on *L. burttii*. To conduct quantitative RT-PCR (qRT-PCR) and transcriptomic sequencing, the root RNAs of *L.*

burttii and *L. japonicus* Gifu were used. The seed bag numbers of the above-mentioned plant materials are listed in Table S2. The information of LORE1 mutant lines is listed in Table S3.

1.3 Plant growth and inoculation

Seeds were scarified using sand paper (grit size 100), surface-sterilized with a sterilization solution (1.2% NaClO and 1% SDS) for 8 min, rinsed with distilled water three times, and soaked in distilled water at room temperature for at least 2 h. Seeds were then germinated on 1/2 B5 medium agar plates for three days at 24°C in dark and three days under a long-day photoperiod (16 h:8 h, light:dark). Six-day-old seedlings were transferred into tulip-shaped Weck jars (Weck 745) containing 300 ml of sterilized growth substrate (sterilization at 180°C for 8 h and cooled down to room temperature before use) supplemented with 40 ml or 90 ml of FAB medium and kept under a long-day photoperiod. After two days, each plant was inoculated with 1 ml of rhizobia suspension (OD_{600} = 0.005) or 1 ml FAB medium as a control. Plants were further kept in long-day photoperiod until the harvest point. After harvest, plants were phenotyped under an MZ16 FA stereomicroscope (Leica). The detailed experimental conditions are described in section 1.4.

1.4 Experimental conditions

1.4.1 Nodulation analysis of *Lotus* species and accessions

The sand was washed with water and dried before use in every case. The autoclaved sand-vermiculite mixture (sand:vermiculite, 1:2) supplied with a 40 ml FAB medium was used to grow plants for phenotyping the nodulation of *Lotus* species and accessions. Seedlings of *L. burttii*, *L. corniculatus*, and eight *L. japonicus* ecotypes were inoculated with *Ml* MAFF, *Rl* Norway, *Ef* HH103 and *Au* LMGT prepared as described in section 1.3. The plants were harvested and phenotyped at 4 weeks post-inoculation (wpi) (*Ml* MAFF) or 6 wpi (*Rl* Norway, *Ef* HH103, and *Au* LMGT).

1.4.2 Examining the effect of growth substrates on nodulation

To study the effect of different substrates on nodulation, experiments were performed with four distinct growth substrates, including sand-vermiculite mixture (1:2), two clay granules (SERAMIS® and Lamstedt), and sand. All the substrates were dried at 60°C for five days before use. The substrates were then used in the preparation of the Weck jars as described in section 1.3. In these experiments, *L. japonicus* Gifu seedlings were grown in Weck jars supplied with either 40 ml (low substrate moisture) or 90 ml (high substrate moisture) FAB medium and kept under a long-day photoperiod. After two days, each plant was inoculated with 1 ml *Rl* Norway as described in section 1.3. Plants were harvested and phenotyped at 6 wpi.

1.4.3 Examining the effect of substrate moisture and nutrient quantity on nodulation

To study the effect of substrate moisture on nodulation, *L. japonicus* Gifu seedlings were grown in Weck jars containing 300 ml sand-vermiculite mixture (1:2) and supplied with either 40, 65, or 90 ml 1X FAB medium together with the inoculum. To examine the impact of nutrient availability on nodulation, plants were supplied with one-time the total nutrient amount or two-times the total nutrient amount in combination with low and high substrate moisture. To control the substrate moisture and nutrient availability, 0.5X, 1X, and 2X FAB mediums were prepared. A total of four conditions were used in the experiment, including a low substrate moisture condition with one-time the nutrients (40 ml of 1X FAB medium), a low substrate moisture condition with two-times the nutrients (40 ml of 2X FAB medium), a high substrate moisture condition with one-time the nutrients (90 ml of 0.5X FAB medium), and a high substrate moisture condition with two-times the nutrients (90 ml of 1X FAB medium). The inoculum was added to the FAB medium accordingly. The plants were kept in long-day photoperiod, harvested, and phenotyped at 6 wpi.

1.4.4 Nodulation analysis of *rlp*, *snf1-1*, and *snf2-2* mutant lines

Seedlings of *rlp1* (L30138037), *rlp2* (L30072908 and L30121643), *rlp3* (L30103870, L30068880, and L30112377), *snf1-1*, and *snf2-2* mutant lines were inoculated with *Rl* Norway as described in section 1.3. A pre-dried sand-vermiculite mixture (60°C treatment for five days before use) supplied with a 40 ml FAB medium or 90 ml FAB medium was used to grow the plants in a long-day photoperiod. The plants were harvested and phenotyped at 6 wpi.

1.4.5 Chemical treatments

To examine the effect of ethylene and naringenin on nodulation, *L. japonicus* Gifu plants were treated with 0.5 µM 1-aminocyclopropane-1-carboxylic acid (ACC, ethylene precursor), 1 µM aminoethoxyvinylglycine (AVG, ethylene biosynthesis inhibitor), or 1 µM naringenin in combination with *Rl* Norway inoculation. Plants were grown in low and high substrate moisture under a long-day photoperiod. The low and high substrate moisture conditions (LM and HM) were defined as 300 ml sand-vermiculite supplied with 40 ml or 90 ml FAB medium, respectively. All the plants were harvested and phenotyped at 6 wpi.

2 DNA extraction and genotyping

2.1 Genomic DNA extraction for *L. burttii*, *L. japonicus* Gifu, and the F1 progeny of a cross between *L. burttii* and *L. japonicus* Gifu

Genomic DNA (gDNA) was extracted from the leaves of *L. burttii*, *L. japonicus* Gifu and the F1 progeny of a cross between *L. burttii* and *L. japonicus* Gifu. The extraction followed previously described methods with few modifications (Murray & Thompson, 1980). In brief, one leaf was taken from each plant at 6 wpi, frozen in liquid nitrogen, and lysed at 30 Hz for 1 min using an MM 400 tissue lyser (Retsch). The tissues were then re-suspended in 500 µl extraction buffer (2% w/v CTAB, 1.42 M NaCl, 20 mM EDTA, 100 mM Tris-HCl, pH8) with 3.1 µl beta-mercaptoethanol and incubated at 65°C for 20 min. Each sample was later mixed with 300 µl of chloroform and centrifuged at 20,000 g for 5 min. Next, the supernatants were mixed with 1/10 volume of 3 M NaOAc and centrifuged at 20,000 g for 15 min to obtain the DNA pellets. The pellets were washed two times with 70% ethanol, air-dried, and resuspended in 50 µl distilled water. The extracted gDNA from *L. burttii* and *L. japonicus* Gifu was used for sequencing and cloning. The extracted gDNA from the F1 progeny of a cross between *L. burttii* and *L. japonicus* Gifu was used for genotyping.

2.2 Genomic DNA extraction for LORE1 lines

The seeds of mutant lines of *RLP1*, *RLP2*, and *RLP3* were germinated and propagated for genotyping. One leaf of each six-day-old plant was used for gDNA extraction via a quick extraction method (Kotchoni & Gachomo, 2009). Samples were frozen in liquid nitrogen and lysed at 30 Hz for 1 min using an MM 400 tissue lyser (Retsch). The lysed tissues were later resuspended in a 400 µl quick extraction buffer (1% SDS 0.5 M NaCl) by vortexing thoroughly for 20 s, followed by centrifugation at 20,000 g for 2 min. Half of the supernatant was transferred into a new tube, mixed with the same amount of isopropanol by inverting gently 10 times, and centrifuged at 20,000 g for 3 min. The obtained pellets were washed two times by 70% ethanol, air-dried, and resuspended in 50 µl distilled water.

2.3 Genotyping of F1 progeny of a cross between *L. burttii* and *L. japonicus* Gifu and LORE1 lines

The gDNA purified as described in the sections above was used for genotyping. The *Lotus* power marker TM1203 and the primers flanking the mutation insertion sites were used for the F1 progeny of a cross between *L. burttii* and *L. japonicus* Gifu and LORE1 lines, respectively (Table S4). PCR was performed using a GoTaq® polymerase in 15 µl reactions using 0.75 units of GoTaq® polymerase, 3 µl buffer, 0.3 µl of 10 mM dNTPs, 0.3 µl of 50 mM $MgCl_2$, 0.3 µl of 10 pmol/µl primers and 10 ng of template DNA. The amplification was conducted under

the following conditions: template was denatured at 95°C for 2 min, then 35 cycles of 95°C for 30 s, 57°C for 30 s and 72°C for 15 s, followed by 72°C for 3 min in the end. The PCR products of LORE1 insertions were analyzed by electrophoresis in a 1% agarose gel at 140 V for 35 min. To obtain a better size resolution, the PCR products of the progeny of *L. burttii* x *L. japonicus* Gifu were analyzed in a 10% acrylamide gel at 120 V for 180 min.

3 Synteny analysis

Syntenic and collinear blocks were identified across 4 selected legume species in the *RLP* cluster by CoGe: SynFind web tool (https://genomevolution.org/coge/SynFind.pl) (Tang et al., 2015). The genomes used to conduct the synteny analysis were *L. japonicus* MG-20 (id 29133), *L. japonicus* Gifu (id 58121), *Medicago truncatula* A17 (id 22583), and *Phaseolus vulgaris* cv. G19833 (id 37644). The 'applied comparison' algorithm and 'specify' feature were Last and Lj2g3v2904830.1, respectively.

4 Genome re-sequencing, SNP identification, and re-annotation

To verify the genomic sequences of four candidates in *L. burttii*, part of the *NFR1*-linked region was amplified by PCR with Phusion® High-Fidelity DNA Polymerase. Amplification reaction volumes were 15 µl using 0.3 unit of Phusion® High-Fidelity DNA Polymerase, 3 µl buffer, 0.3 µl of 10 mM dNTPs, 0.3 µl of 50 mM $MgCl_2$, 0.3 µl of 10 pmol/µl primer and 10 ng of template DNA. The amplification was conducted under the following conditions: template was denatured at 98°C for 2 min, then 35 cycles of 98°C for 30 s, 60°C for 30 s and 72°C for 5 min, followed by the last step at 72°C for 5 min. The PCR products were analysed by gel electrophoresis and purified by GeneJET Gel Extraction Kit (K0692, Thermo Fisher Scientific Inc., USA), sequenced by Sanger sequencing (Sequencing Service, LMU Faculty of Biology). Primers used for amplification and sequencing are listed in Table S4.

Sequences of *L. burttii* were compared against *L. japonicus* Gifu genome v1.2 (Kamal et al., 2020) using the integrative genomic viewer (IGV 2.8.3) (Robinson et al., 2011). Single nucleotide polymorphisms (SNPs) between both species were identified and annotated in CLC Main Work Bench 20 (QIAGEN). To validate the automatic annotation of the *L. japonicus* Gifu genome, the gene structure in the genomic region between LotjaGi1g1v0060000 (nucleotide position: 8360608 to 8369978) and LotjaGi1g1v0062900 (nucleotide position: 8755286 to 8763613) was manually predicted by web tool FGENESH (http://www.softberry.com/berry.phtml) (Solovyev et al., 2006) and then annotated by NCBI blastx (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

5 Plasmid construction for trans-complementation assay

5.1 Golden Gate cloning

The Golden Gate cloning method was used in this study to create and assemble different cloning constructs. The constructs were categorized into LI, LII, and LIII as described in (Binder et al., 2014). T4 DNA ligase (M0202S, New England Biolabs Inc., USA) and specific type II restriction enzymes (BpiI, BsaI and Esp3I) were used in a one-step cut-ligation reaction to assemble desired cloning elements. For LI and LIII constructs, BpiI (ER1011, Thermo Fisher Scientific Inc., USA) was used. For LII constructs, BsaI (R0535S, New England Biolabs Inc., USA) was used. For replacing the element at the promoter position, Esp3I (R0734S, New England Biolabs Inc., USA) was used. Cut-ligation reaction volumes were 15 µl using 1.5 µl 10X restriction enzyme buffer, 7.5 units type II restriction enzyme, 300 units T4 ligase, 1.5 µl 10 mM ATP and DNA fragments. For constructing LI plasmids, a molecular weight ratio insert:vector of 3:1 was used. For LII and LIII constructs, 100 ng of every element were used. Cut-ligation reactions were conducted in a thermocycler under the following conditions: 10 cycles between 37°C 10 min and 16°C 10 min, and followed by a final ligation step at 16°C overnight. The constructs generated and used in this section are listed in Table S5.

5.2 Construction of LI *LbRLP4* and LI *LbRLP4$_{pro}$*

To obtain the LI construct of *RLP4* of *L. burttii* (*LbRLP4*) as well as the promoter sequence of *RLP4* (*LbRLP4$_{pro}$*), the genomic sequences of *LbRLP4* and *LbRLP4$_{pro}$* were amplified from *L. burttii* gDNA. The *LbRLP4* sequence was divided into 4 fragments using primers designed to introduce silent mutations, which remove the endogenous BpiI and BsaI recognition sites. Additional BpiI and BsaI recognition sites required for cloning were introduced in the 5'-end of the primer (Table S4, cloning primers). The fragments of *LbRLP4* and *LbRLP4$_{pro}$* were amplified by PCR with Phusion® High-Fidelity DNA Polymerase following the manufacturer's instructions. The amplification was conducted under the following conditions: template was denatured at 98°C for 2 min, then 35 cycles of 98°C for 30 s, 63°C for 40 s, and 72°C for 3 min, followed by a final step at 72°C for 3 min. The amplified gene and promoter fragments were purified using GeneJET Gel Extraction Kit (Thermo Fisher Scientific Inc., USA) and then assembled by BpiI cut-ligation into a LI BB3 vector, respectively. Two LI constructs were generated. One plasmid contained 2975 nt of *LbRLP4* genomic sequences from the start codon (ATG) to the end of coding sequence without the stop codon and the other plasmid contained a 602 nt of *LbRLP4$_{pro}$* genomic sequences upstream the start codon.

5.3 Construction of LII expression plasmids

5.3.1 LII transformation marker

A LII construct to drive the constitutive expression of an mCherry fluorescent protein was designed and used as a transformation marker *in planta*. For this, Golden Gate cloning elements including the $35S_{pro}$, a nuclear localization signal (NLS), an element encoding mCherry fluorescent protein, a heat shock protein terminator and a dummy element were assembled into a LII expression vector (BB20) via BsaI cut-ligation.

5.3.2 Expression constructs of *LbRLP4*

The LII expression constructs of *LbRLP4* were designed to be driven by either a strong constitutive promoter or its native promoter. The *LbRLP4* over-expression construct consisted of elements including the *L. japonicus Ubiquitin1* promoter ($LjUbq1_{pro}$), *LbRLP4* gene, a 6x His tag, the 35S terminator, and two dummies. To construct the native promoter-driven *LbRLP4* expression plasmid, the $LjUbq1_{pro}$ was replaced by $LjRLP4_{pro}$. All elements were assembled into a BB26 LII expression vector via BsaI cut-ligation.

For trans-complementation assays of *LbRLP4*, LIII plasmids containing expression cassettes of the transformation marker and either over-expression or native promoter-driven *LbRLP4* were generated. Together with dummy elements, the expression cassettes of transformation marker and *LbRLP4* were assembled via BpiI cut-ligation into a BB52 LIII backbone.

5.3.3 Expression constructs of *LbRLP2*

LIII constructs containing an expression cassette of the transformation marker and an expression cassette of either over-expression or native promoter-driven *LbRLP2* were generated for the trans-complementation assays of *LbRLP2*. In detail, a LII *LbRLP2* intermediate plasmid comprising a *LacZ* element at the promoter position, a *LbRLP2* gene and a Nos terminator was generated. This *LbRLP2* intermediate plasmid together with dummy elements and the expression cassette of the transformation marker were then assembled via cut-ligation mediated by T4 DNA ligase and BpiI into a BB52 LIII backbone. The *LacZ* element was later replaced by either the *Arabidopsis thaliana Ubquitin10* promoter ($AtUbq10_{pro}$) or the $LbRLP2_{pro}$ via Esp3I cut-ligation.

5.3.4 Expression constructs of *LbRLP2+4*

For the *LbRLP2* and *LbRLP4* trans-complementation assays, LIII constructs containing expression cassettes of the transformation marker, *LbRLP2* and *LbRLP4* were generated. In the LIII constructs, *LbRLP2* and *LbRLP4* were either driven by their native promoters or the

over-expression promoters. The expression cassettes of the transformation marker, the *LbRLP2* intermediate plasmid and the expression plasmid of *LbRLP4* were assembled via BpiI cut-ligation. The *LacZ* element of the *LbRLP2* intermediate plasmid was replaced by either the $AtUbq10_{pro}$ or the $LbRLP2_{pro}$ via Esp3I cut-ligation.

5.4 Plasmid extraction and transformation

5.4.1 Plasmid selection

The obtained cut-ligation products generated as described in the above sections were transformed into chemocompetent *E. coli* TOP10 cells by heat-shock method (Hanahan, 1983) for selection and proliferation. For LI and LII constructs, 5 µl of each cut-ligation product were gently mixed with 50 µl *E. coli* TOP10. In the case of LIII constructs, 10 µl of each cut-ligation product were used for transformation. The mixtures were then incubated on ice for 30 min, followed by a heat-shock at 42°C for 45 s, and an immediate incubation on ice for 2 min. Then, 1 ml LB medium was added and the bacteria were incubated for 1 h at 37°C and 450 rpm to ensure regeneration. Afterward, 100 µl of each bacterial suspension were spread on LB agar plate supplemented with appropriate antibiotics (25 ng μl^{-1} Gentamycin for LI and LIII plasmid selection; 100 ng μl^{-1} Spectinomycin for LII plasmid selection) at 37°C overnight.

5.4.2 Plasmid purification

To purify the plasmids from the transformed *E. coli* TOP10 strains, a single colony of each construct was picked and inoculated into 3 ml LB broth supplemented with the appropriate antibiotics and grown at 37°C and 180 rpm for 12-16 h. Then, 2 ml of the bacteria culture were centrifuged for 2 min at 6,800 g at room temperature. After discarding the supernatant, the pellet was resuspended in 190 µl P1 buffer (50 mM Tris-base, 10 mM $Na_2EDTA{\bullet}2H_2O$, pH 8.0) supplemented with 10 µl RNase A (10 µg μl^{-1}, Sigma-Aldrich). Next, 200 µl of P2 buffer (0.2 M NaOH and 1% SDS) were added, gently mixed and the mixtures were then incubated at room temperature for 5 min. Then, 200 µl of P3 buffer (3 M potassium acetate, pH 5.5) were added and mixed gently, followed by centrifugation for 10 min at 13,000 g. The supernatant was mixed with pre-cooled isopropanol in a one-to-one ratio and centrifuged for 10 min at 13,000 g. The obtained pellet was then washed twice with 500 µl of pre-cooled 70% ethanol, centrifuged for 3 min at 13,000 g and dried at 60°C for 5-10 min in a Thermomixer comfort (Eppendorf). The DNA pellet was resuspended in 50 µl distilled water and the plasmid concentration was estimated using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies Inc.).

5.5 Quality control of cloning

After purification, the acquired LI constructs with target sequences were confirmed via sequencing with M13F (5'-TGTAAAACGACGGCCAGT-3') and/or M13R (5'-GGAAACAGCTATGACCAT-3') primers (Sequencing Service, LMU Faculty of Biology) as well as a digestion-pattern control via BsaI restriction enzyme. A digestion-pattern control was applied to the LII and LIII constructs with BpiI and BsaI restriction enzymes, respectively. The digestion products were analyzed by electrophoresis in a 1% agarose gel at 140 V for 35 min.

6 Trans-complementation assay

6.1 *Agrobacteria* transformation

For every target plasmid, 100 ng of the plasmid DNA were mixed with 50 µl of *A. rhizogenes* strain AR1193 in an electroporation cuvette. The target plasmids were introduced into *Agrobacteria* via electroporation, using EcI electroporation program of a MicroPlulser™ (BIO-RAD). After electroporation, 1 ml LB was added to the *Agrobacteria*, and the bacteria suspension was incubated at 28°C and 180 rpm for 2 h. Next, 10 µl of *Agrobacteria* suspension of each construct were diluted 10 times in YEB medium and then spread on YEB agar plates supplemented with 50 µg ml^{-1} of carbenicillin, rifampicin, and kanamycin. Plates were incubated at 28°C for 2 nights. Every *Agrobacterium* culture was freshly prepared as follows: 3 ml liquid pre-cultures were made in YEB from the single colonies and incubated at 28°C for 2 nights. From every liquid culture, 100 µl were taken and spread on LB agar plates followed by overnight incubation at 28°C. This overnight culture of the *Agrobacteria* was resuspended in sterile water for transformation. The plasmids used in the trans-complementation assays were listed in Table S5.

6.2 Hairy root transformation

To express *LbRLP2* and/or *LbRLP4* transiently in *L. japonicus* Gifu, trans-complementation assays via *Agrobacteria*-mediated hairy root transformation were performed as described previously (Díaz, 2005). Seeds of *L. japonicus* Gifu were germinated following the method described in section 1.3. In addition, *A. rhizogenes* AR1193 bacteria carrying the *LbRLP2* and/or *LbRLP4* expression plasmids were activated from glycerol stocks on selective LB agar plates and cultured at 28°C for 2 nights.

Six-day-old seedlings were placed on a petri dish and the roots were soaked in 1 ml *Agrobacteria* suspension. The roots of the seedlings were removed and the remaining hypocotyls were dipped into *Agrobacteria*. Transformed plants were kept at 24°C in the dark

for 2 days and transferred to a long-day photoperiod. After 3 days, plants were transferred onto B5 medium containing cefotaxime (300 µg ml^{-1}) to clear the *Agrobacteria*. After 21 days, plants were screened for mCherry fluorescence. Transformed plants were transferred into sterilized Weck jars filled with sand-vermiculite mixture (1:2) and 40 ml FAB medium. After two days, each plant was inoculated with 1 ml *Rl* Norway suspension (OD_{600} = 0.005). Plants were kept in a long-day photoperiod and harvested 8 wpi.

The roots of harvested plants were cleaned with water and fixed with 4% formaldehyde in 50 mM PIPES buffer (pH 7.0) via vacuum infiltration for 20 min followed by another 10-min vacuum infiltration. The roots were incubated at room temperature for 45 min and then washed three times with 50 mM PIPES buffer (pH 7.0) for 5 min. After washing, the roots were observed under MZ16 FA stereomicroscope (Leica). The nodulation events were counted and recorded. The representative nodulation phenotypes were photographed.

7 CRISPR-Cas12a genome editing assay on *L. burttii*

7.1 Guide RNA (gRNA) design

A CRISPR-Cas12a genome editing method was performed on *L. burttii* for creating mutants in genes of interest following an established protocol (M. Bircheneder, unpublished). The gRNAs were designed to target *LbRLP1*, *LbRLP2*, *LbRLP3*, and *LbRLP4* using the web tool CRISPOR (Concordet & Haeussler, 2018). The *L. japonicus* genome was chosen as reference and the program 'Cas12a (Cpf1)-21bp guides-recommend by IDT' was chosen for protospacer adjacent motif prediction. The predicted gRNAs were selected with high predicted efficiency and low off-targets. Sequences of the chosen gRNAs are listed in Table S4.

7.2 Plasmid construction

To make the sequences compatible for Golden Gate cloning, additional BpiI and BsaI recognition sites were added at the 5'- and 3'-ends of the gRNA oligos. The designed oligos and their reverse complementary sequences were ordered from Sigma Aldrich and resuspended in distilled water to 100 µM. To anneal the pairing oligos, 5 µl were taken from each 100 µM oligo suspension, incubated at 98°C for 5 min, and cooled down at room temperature. The paired oligos were cloned into BB3 LI vector via BpiI cut-ligation.

The acquired LI constructs contained the target gRNA and the following cloning elements: the $LjU6_{pro}$, the *LjU6* terminator (*LjU6*-T), a dummy, and the BB26 LII vector. Elements were assembled into LII constructs via a BsaI cut-ligation. The LIII constructs were generated by

assembling three LII constructs, including a hygromycin resistance gene driven by Nos_{pro}, a *Cas12a* driven by $LjUbq1_{pro}$ and gRNA driven by $LjU6_{pro}$.

The cut-ligations were performed following the method described in section 5.1. All cut-ligation products were transformed into *E. coli* TOP10 for selection and plasmid extraction, as described in sections 5.4 and 5.5. The oligos for cloning are listed in Table S4. The constructs used for genome editing are listed in Table S5. The LI $LjU6_{pro}$ and *LjU6*-T elements as well as LII constructs expressing hygromycin resistant gene and *Cas12a* were provided by M. Bircheneder.

7.3 *Agrobacteria* transformation

The target constructs were introduced into *A. tumefaciens* Agl1 via electroporation, using Agl electroporation program of a MicroPlulser™ (BIO-RAD). The electroporation was conducted as described in section 6.1. The transformed *A. tumefaciens* Agl1 were cultured on YEB agar plates and incubated at 28°C for 2 nights.

7.4 Stable transformation and tissue culture

7.4.1 Seed germination and hypocotyl transformation

L. burttii seeds were geminated with the method described in section 2.1 with few modifications. After sterilization, the seeds were placed in square Petri dishes with 1 layer of 1.5 mm thick (Ahlstrom Munksjö) and 2 layers of 0.35 mm thick (Hahnemühle) sterile blotting paper (11 cm x 11 cm) soaked in 50 ml sterile water. The plates were kept at 24°C in the dark for 3 days and transferred to a long-day photoperiod for 3 days.

Six-day-old *L. burttii* seedlings and the transformed *A. tumefaciens* Agl1 were prepared for generating mutant lines for the genes of interest. The *Agrobacteria* on YEB plates were suspended in a 4 ml YMB/potassium phosphate buffer mixture (Table S6, S7) and added to a sterile 9 cm diameter Petri dish with a thin layer of Whatman filter paper, on which the *L. burttii* seedlings were later placed. The shoots and roots of *L. burttii* were firstly removed from hypocotyls, and the remaining hypocotyls were cut into pieces of 3 to 4 mm. The tissues were always immersed in the *Agrobacteria* suspension. The transformed hypocotyls were then placed on the freshly prepared co-cultivation medium (Table S7) with 2 layers of 0.35 mm and 1 layer of 1.5 mm sterile blotting paper soaked in 50 ml sterile water. The hypocotyls and *Agrobacteria* were co-cultured in the dark for 6 days at 21°C.

7.4.2 Callus induction

After 6 days of co-cultivation, the thin piece of Whatman filter paper with the transformed hypocotyls on top was transferred into a square Petri dish with callus induction medium containing hygromycin B (Table S6 and S7) for selecting the transformed tissues and incubated at 24°C under a long-day photoperiod. After 1 week, the transformed hypocotyls were transferred to a freshly prepared callus induction medium. This step was repeated every 7 days for 5 weeks and kept at 24°C under a long-day photoperiod. The emerging calli were separated from the hypocotyls, which turned brown, and were kept in a callus induction medium.

7.4.3 Shoot induction and elongation

The growing calli were later transferred to shoot induction medium and incubated in a long-day photoperiod at 24°C for 7 days. The shoot induction medium was exchanged every 7 days for 2 - 7 weeks based on the development of calli.

The calli became a deeper green color after a few weeks of growing in the shoot induction medium. The leaf-like shoot primordia then appeared on calli, which were transferred to shoot elongation medium and incubated under a long-day photoperiod at 24°C for 7 days. Fresh medium was applied every 7 days for 3 to 6 weeks until the shoot primordia became circa 1 cm-long. The composition of each media is listed in Tables S6 and S7.

7.4.4 Root induction and elongation

Emerging shoots that were circa 1 cm in length with leaves were cut and transferred to root induction medium and incubated for 10 days under long-day photoperiod at 24°C. After 10 days, shoots with emerging roots were vertically placed in magenta boxes with 100 ml of root elongation medium and grown under a long-day photoperiod at 24°C for 3 to 4 weeks. The composition of the media is listed in Tables S6 and S7.

7.5 Plant growth and reproduction

Plants with approximately 2 cm long roots were transferred to pots with Stender substrate (A210, Stender GmbH) fertilized with Osmocote Exact Standard fertilizer (3 g fertilizer per litter substrate) and kept in a greenhouse for reproduction. The light intensity was 400 W. Additional lighting from 6 to 10 am and 3 to 10 pm was provided. The day and night temperatures were 21-24°C and 18-21°C, respectively.

7.6 DNA extraction and genotyping

Ten plants of each transformed construct were chosen for genotyping. One leaf from each plant was taken for gDNA purification with the quick extraction method described in section 2.2. The genotyping was performed by amplifying the region flanking the location of gRNA in the gene of interest. The PCR was conducted with Phusion® polymerase in 50 µl reaction following the manufacturer's instructions with a modification that only 0.1 µl Phusion® polymerase was used per reaction. The amplification was conducted under the following conditions: template was denatured at 98°C for 1 min, then 35 cycles of 98°C for 30 s, 63°C for 30 s, and 72°C for 7 min, followed by a final step at 72°C for 7 min at the end. The products were analyzed by electrophoresis in 1% agarose gel at 140 V for 35 min. The products were purified by GeneJET Gel Extraction Kit (Thermo Fisher Scientific Inc., USA) following the manufacturer's instructions and sequenced by Sanger sequencing (Sequencing Service, LMU Faculty of Biology). The primers used for genotyping are listed in Table S4.

8 Transcriptome sequencing (RNA-seq) and data analysis

8.1 Plant growth conditions

L. burttii and *L. japonicus* Gifu seedlings were prepared as described in section 1.3. Six-day-old *L. burttii* and *L. japonicus* Gifu seedlings were grown in LM and inoculated either with 1 ml *RI* Norway or FAB medium as a control. Additionally, *L. japonicus* Gifu seedlings were grown in HM and inoculated either with 1 ml *RI* Norway or FAB medium as a control. Whole root systems were collected from treated plants at 2 wpi for RNA extraction.

8.2 RNA isolation

For every condition, three roots were pooled together as one biological replicate. Six replicates were used for the *RI* Norway inoculated condition of both plant species and the mock condition of *L. burttii*; 9 replicates were used for mock treated *L. japonicus* Gifu. Samples were frozen in liquid nitrogen and then lysed by an MM400 tissue lyser (Retsch). Total RNA was extracted by Spectrum™ Plant Total RNA-Kit (STRN250, Sigma-Aldrich, USA) according to the manufacturer's instructions. The extracted RNA was treated with DNaseI (Ambion Inc., USA) to eliminate DNA. The RNA quality and concentration were measured by a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies Inc., USA).

8.3 cDNA library preparation

The cDNA library was prepared following the methods described previously with modifications (Janjic et al., 2021) (in collaboration with L. Wange and W. Enard). Prior to library preparation,

RNA concentrations were normalized to 1 ng μl^{-1}. 4 ng of RNA from each sample, together with 5 µl reverse transcription mix (Maxima H Minus reverse transcriptase 0.15 µl, Maxima RT 5X buffer 2 µl, 25 mM dNTP 0.4 µl, 100 µM TSO 0.1 µl and UltraPure water 2.35 µl) and 1 µl barcoded oligo dT (10 µM) were incubated at 42°C for 90 min. Subsequently, all samples were pooled into one reaction and cleaned up using homemade SPRI beads (Sera Mag SpeedBeads in 22% PEG, 1 M NaCl, 0.01 M Tris-HCl, $1x10^{-3}$ M EDTA, 0.01% IGEPAL and 0.05% Sodium Azide). Beads were applied to the pooled cDNAs at a 1 to 1 ratio with 5 min incubation, followed by washing twice with 1 ml 80% EtOH followed by air drying the beads. The cDNAs were eluted in 17 µl UltraPure water. Purified cDNAs were treated with exonuclease I at 37°C for 20 min followed by 80°C for 10 min to inactivate the enzyme. The cDNAs were purified again by homemade SPRI beads with steps mentioned above and finally eluted in 20 µl UltraPure water.

The full-length cDNAs were mixed with amplification mix (25 µl KAPA HiFi 2X RM, 3 µl 10 µM pre-amp primer and 2 µl UltraPure water) and amplified by PCR in a thermocycler under the following conditions: 98°C 3 min, 10 cycles of 98°C 15 s, 65°C 30 s, and 72°C 4 min, followed by 72°C 10 min for a final extension. The amplified cDNAs were purified by homemade SPRI beads. The quality and quantity of the cDNAs were later qualified by Agilent 2100 Bioanalyzer with High Sensitivity DNA Analysis Kit and quantified by Quant-iTPicoGreen dsDNA Assay kit. After passing the quality and quantity control the pooled cDNA libraries were used to construct sequencing libraries with Nextera XT Library Prep Kit. Three replicates with 0.8 ng cDNA each were tagmented according to the manufacturer's protocol and PCR amplified using a 3'-specific primer to capture the barcode and UMI sequences introduced during the reverse transcription.

8.4 Transcriptomic sequencing and data processing

Libraries were paired-end sequenced on an Illumina Hiseq 1500 instrument, with a 16 bp long first read covering the barcode and UMI sequences and 50 bp long second read covering the cDNA sequences. Samples were sequenced to a depth of 5-10 Mio raw reads per sample (LAFUGA Gene Center, LMU Munich). The raw data was pre-processed, mapped and a count matrix was generated using the zUMIs pipeline (2.5.4) (Parekh et al., 2018) with STAR (2.6.0) (Dobin et al., 2013). Reads were mapped to *L. japonicus* Gifu v1.2 genomes with the corresponding gene annotation (in collaboration with L. Wange and W. Enard).

8.5 Gene expression analysis

8.5.1 Untargeted approach

The comparisons of the expression profiles between conditions were performed using the DESeq2 package (Love et al., 2014) in R (version 3.6.1) and the variances between conditions were assessed by plotPCA function in the DESeq2 package. In brief, the genes with total read counts over 1 were taken and normalized and the apeglm estimation (Jarvis et al., 1982) within the lfcShrink function were called and applied by DEseq2 package. Differentially expressed genes (DEGs) were identified within plant species (e.g. *L. japonicus* Gifu LM vs HM) and within conditions (eg. *L. burttii* vs *L. japonicus* Gifu in LM). The criteria for DEGs comparisons were $\alpha = 0.01$, $p < 0.01$. The obtained DEGs were filtered by absolute value of $\log_2$ fold change >1 and the numbers of DEGs in each comparison were displayed using Venn Diagrams.

Gene ontology (GO) analysis was applied to the DEGs, performed in R version 3.6.1 with the topGO package (Alexa & Rahnenfuhrer, 2019) using the Fisher statistic method and the classic algorithm. The file for GO-terms of *L. japonicus* Gifu (*L. japonicus* Gifu v1.2 genome) was obtained from Lotus Base (https://lotus.au.dk/) (Mun et al., 2016).

8.5.2 Targeted approach

To investigate the expression of individual genes involved in symbiosis, hormone signaling, defense, and flavonoid biosynthesis, the gene ID of each gene of interest was obtained from open sources (NCBI: https://www.ncbi.nlm.nih.gov/ and Lotus Base: https://lotus.au.dk/) by using respective proteins as search queries. The acquired gene ID was used to extract the read counts of each gene from the RNA-seq data in R version 3.6.1. The R packages, including dplyr (Wickham et al., 2021a), readr (Wickham & Hester, 2020), vctrs (Wickham et al., 2021b), and stringr (Wickham, 2019), were used for data sorting.

8.6 Network analysis

Genes differentially expressed between low and high humidity conditions with absolute value of $\log_2$ fold change >1 and *p*-value <0.1 were selected for network analysis. The expression data of *L. japonicus* Gifu v1.2 genome were retrieved with DEGs from 6 conditions, including leaf, mature flower, seed, root, and nodule at 10 and 21 dpi (Lotus Base, https://lotus.au.dk/expat/). The co-expression network was built using the WGCNA package (Langfelder & Horvath, 2008) in R version 3.6.1. First, an adjacency matrix was built based on Pearson's correlations between each gene and further used in topological overlap measurements (TOM) and corresponding dissimilarity (diss-TOM) calculations. The diss-TOM was used to distinguish the distance between hierarchical gene clusters. The clusters of highly

interconnected genes were identified by the DynamicTree Cut algorithm (Langfelder et al., 2008) with the minimum number of genes per module set to 50 and the standard threshold set to 0.25. The results were plotted with the function plotDendroAndColors.

9 qPCR validation of gene expression

SuperScript™ III reverse transcriptase (Invitrogen, USA) was used to synthesize cDNA with 270 ng total RNA of roots extracted from section 8.2 after DNaseI treatment following the manufacturer's instructions. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) volumes were 7 µl using 2 µl 1:10 (v/v) diluted cDNA, 0.3 µM of each primer, and 3.5 µl 2x Fast SYBR™ Green Master Mix (Thermo Fisher Scientific Inc., USA) according to the manufacturer's instructions in a total volume of 7 µl. The qRT-PCR was conducted in a 384-well plate with Quantstudio5 system (Thermo Fisher Scientific Inc., USA) in a thermocycler under the following conditions: 95°C for 2 min, 40 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 20 s, followed by dissociation curve analysis. Five to ten biological replicates and 3 technical replicates were performed. Relative expression was normalized to the internal control gene (*ATPsynthase*) and presented as $2^{-\Delta CT}$. The primers used to detect the expression of *RLP1*, *RLP2*, *RLP3*, *RLP4*, and *NIN* are listed in Table S4.

10 Quantification of the root colonization of *RI* Norway

10.1 Quantification of *RI* Norway colonization by colony-forming unit

To quantify the root colonization by *RI* Norway under different conditions, colony-forming unit (CFU) assay was conducted on isolated root fragments. Samples were incubated overnight (approximately 15 h) at room temperature in 500 µl sterile water. Next, the samples were bath sonicated at 50 kHz for 15 min, followed by 2 min of vortexing. Serial dilutions of 10^{-3}, 10^{-5}, 10^{-7} were prepared from the bacteria suspensions. A 150 µl aliquot of each dilution and the original bacteria suspension were plated on TY medium plates supplemented with tetracycline (2 µg ml^{-1}) and streptomycin (500 µg ml^{-1}). The number of colonies was counted after three days of incubation at 28°C, and plates were imaged with a Epson Perfection V700 Photo scanner (Epson).

10.2 Evaluation of *RI* Norway colonization by qPCR

To evaluate the colonization of *RI* Norway under LM and HM conditions, the gDNA of *RI* Norway was extracted from root fragments at 6 wpi. The root tissues were collected, frozen by liquid nitrogen and lysed by an MM 400 tissue lyser (Retsch) at 30 min^{-1} for 30 s twice. The

extractions were performed following a CTAB buffer-mediated extraction protocol (William et al., 2012). The acquired DNA pellets were air-dried at room temperature and re-suspended in 50 µl sterilized water. The qPCR reactions to detect *RI* Norway were set up in a total volume of 7 µl containing 2 µl gDNA, 0.3 µM of each primer and 3.5 µl 2x Fast SYBR™ Green Master Mix (Thermo Fisher Scientific Inc., USA). The amplification was performed using a real-time QuantStudio™ 5 real-time 384-well PCR system (Thermo Fisher Scientific Inc., USA) in a thermocycler under the following conditions: 95°C for 2 min, 40 cycles of 95°C for 30 s, 57°C for 30 s and 72°C for 20 s, followed by dissociation curve analysis. Two to five biological replicates and 3 technical replicates were conducted. The used primers are listed in Table S4.

A standard curve for *RI* Norway quantification was prepared as a reference for quantification. The 10-fold serial dilutions from 10^{-1} to 10^{-8} of a known concentration of *RI* Norway gDNA were determined by Qubit fluorometer and used to produce the standard curve. As the *RI* Norway genome consists of 7,788,085 bp (Liang et al., 2018) and 1 kb of DNA equal approximately 10^{-6} pg (Sessions, 2013), 1 ng of gDNA represents approximately 1.28 x 10^5 cells of *RI* Norway. The amount of *RI* Norway genome was calculated by comparing CT values of the tested samples to the CT values of the standard curve, then converted into the cell number of *RI* Norway.

11 Statistical analysis

All statistical analyses were conducted in R version 3.6.1 (2019-07-05) and the package multcompView (Spencer Graves et al., 2019). The Shapiro-Wilk test was used to examine if the distributions of the numbers of nodule and nodule primordia were normal. ANOVA followed by Tukey's post hoc test, Kruskal-Wallis test followed by Dunn's post-hoc-test or Mann-Whitney U test, and unpaired t-test were performed to show the significant difference between tested groups. The results of the statistical analysis are stated in the figures and shown in lowercase letters. For RNA-seq data analysis, the following packages 'DESeq2' (Love et al., 2014), 'topGo' (Alexa & Rahnenfuhrer, 2019) and 'WGCNA' (Langfelder & Horvath, 2008) were used to perform DEG analysis, GO analysis, and network analysis, respectively. The data plots were generated by Rstudio version 1.1.463 with package 'ggplot2' (Wickham, 2016).

Results

1 Sub-compatible symbiosis partners exhibit variation in nodulation phenotype

Rl Norway induces various nodulation phenotypes in *Lotus* species and ecotypes without fixing nitrogen (Gossmann, 2012). Therefore, *Rl* Norway is considered a sub-compatible symbiosis partner for *Lotus* species. To investigate whether such variation in nodulation phenotypes is strain-specific, ten *Lotus* accessions in combination with four rhizobia strains from different genera were inspected for their nodulation phenotype. In addition to *Rl* Norway, two sub-compatible strains (*Ef* HH103 and *Au* LMGT) and one compatible strain of *Lotus* (*Ml* MAFF) were included.

At 4 wpi, *L. burttii*, *L. corniculatus*, and eight *L. japonicus* ecotypes inoculated with the compatible strain *Ml* MAFF developed pink nodules. In contrast, sub-compatible strains (*Rl* Norway, *Ef* HH103 and *Au* LMGT) induced either nodule primordia, white nodules or no nodulation on the roots of *L. japonicus* accessions after six weeks of inoculation (Figure 4). *L. burttii* inoculated with *Ef* HH103 showed pink nodules (Figure 4). Notably, leaves of *L. burttii* inoculated with *Ml* MAFF and *Ef* HH103 were green, although the shoot length of *L. burttii* inoculated *Ml* MAFF was higher. This finding suggested that sub-compatible symbiotic partnerships possessed a greater variation in symbiotic compatibility than compatible symbiosis.

L. burttii, *L. corniculatus*, *L. japonicus* MG-70, and *L. japonicus* MG-20 nodulated with every tested strain at a high nodulation rate (Figure 4) with variation in the number of nodules and primordia (Figure 5). Although *L. japonicus* MG-136 and *L. japonicus* MG-119 also nodulated with every tested strain, they nodulated with *Au* LMGT at a lower rate, exhibiting a lower number of nodules and primordia (Figure 4, 5). While other *L. japonicus* ecotypes (MG-135, MG-123, and MG-86) nodulated with *Rl* Norway and *Ef* HH103, *L. japonicus* Gifu nodulated only with *Ef* HH103, showing the highest selectivity among the ten selected *Lotus* accessions. Considering the nodulation rate and the number of nodules and primordia, the most striking difference was between *L. burttii* and *L. japonicus* Gifu, the most and least restrictive accession, respectively.

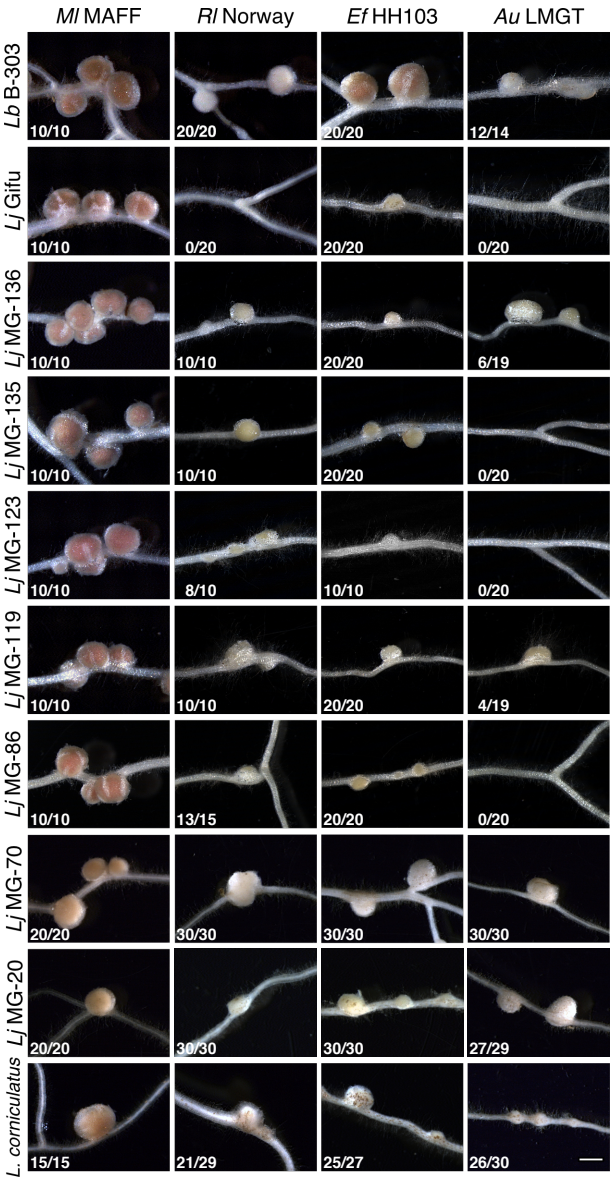


Figure 4. Variation in nodulation phenotype of *Lotus* accessions inoculated with rhizobia strains. Representative nodulation phenotypes of ten selected *Lotus* accessions inoculated with *Mesorhizobium loti* MAFF 303099 (*Ml* MAFF), *Rhizobium leguminosarum* Norway (*Rl* Norway), *Ensifer fredii* HH103 (*Ef* HH103), and *Allorhizobium undicola* LMGT (*Au* LMGT) (OD_{600} = 0.005). Plants inoculated with *Ml* MAFF were observed at 4 wpi, while plants inoculated with *Rl* Norway, *Ef* HH103, and *Au* LMGT were observed at 6 wpi. The numbers on the lower-left corner indicate the number of plants showing nodules and nodule primordia over the total observed plants. Scale bar = 1 mm.

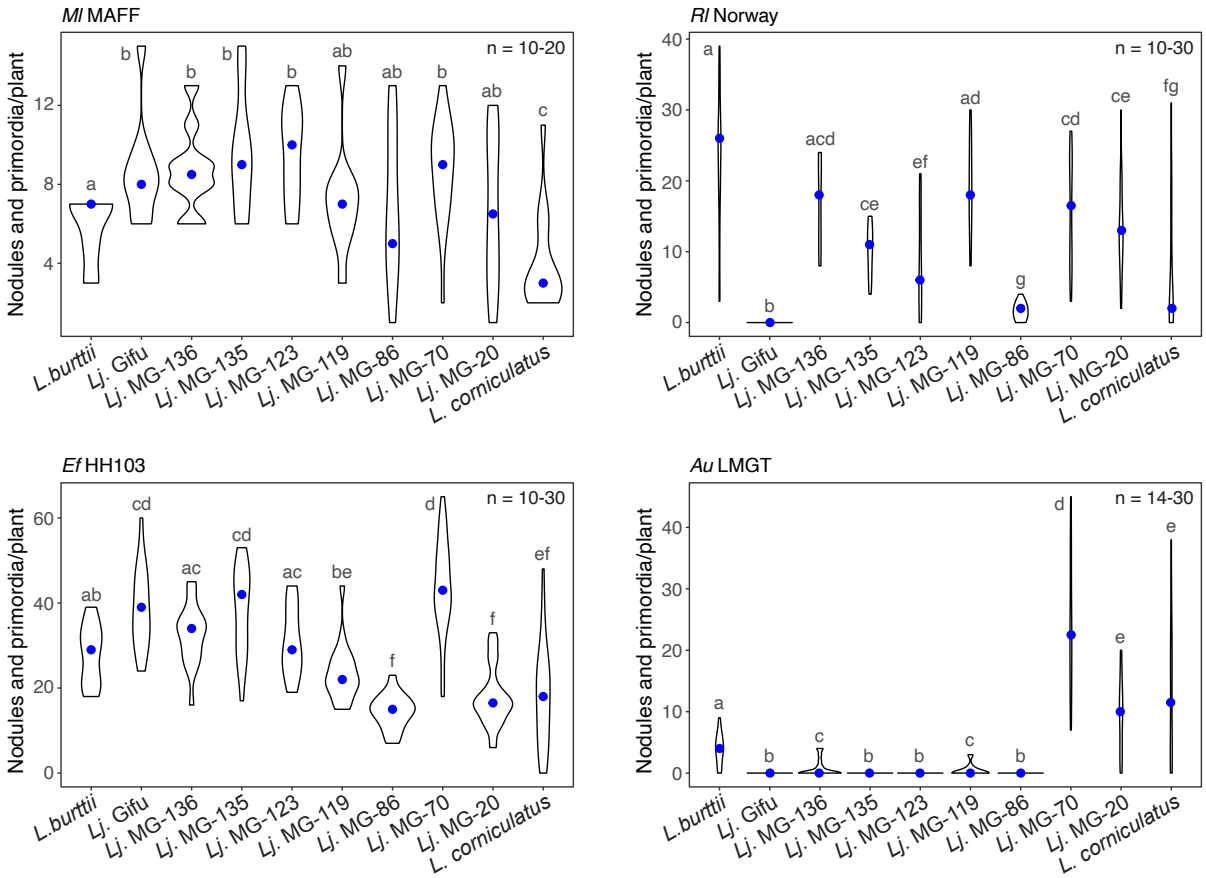


Figure 5. Variation in numbers of nodule and nodule primordia in *Lotus* accessions upon rhizobia inoculation. The number of nodules and primordia of *L. burttii*, *L. corniculatus*, and eight *L. japonicus* ecotypes (Gifu, MG-136, MG-135, MG-123, MG-119, MG-86, MG-70, and MG-20) inoculated with *M. loti* MAFF 303099 (*Ml* MAFF), *R. leguminosarum* Norway (*Rl* Norway), *E. fredii* HH103 (*Ef* HH103), and *A. undicola* LMGT (*Au* LMGT) (OD_{600} = 0.005) is presented. Plants inoculated with *M. loti* MAFF 303099 were observed at 4 wpi, while plants inoculated with *R. leguminosarum* Norway, *E. fredii* HH103, and *A. undicola* LMGT were observed at 6 wpi. Ten to thirty plants of each *Lotus* accession were observed in every inoculation condition. Blue dots indicate the median in each condition. Different lowercase letters indicate significant differences determined by the Kruskal-Wallis test followed by the Mann-Whitney U test (p <0.05).

2 Substrate moisture contributes to nodule formation and rhizobia colonization

2.1 *L. japonicus* Gifu nodulate with *Rl* Norway in high substrate moisture

Although under standard laboratory conditions *L. japonicus* Gifu does not nodulate with *Rl* Norway, occasionally nodules were observed on *L. japonicus* Gifu when the substrate was moister. To define the nodulation conditions for *L. japonicus* with *Rl* Norway, four growth substrates combined with two amounts of FAB medium (50 or 100 ml) were used to grow *L. japonicus* Gifu plants inoculated with *Rl* Norway. At 6 wpi, nodules and primordia developed

on *L. japonicus* Gifu grown in two kinds of clay granulate (Seramis® and Lamstedt) and sand regardless of the amount of FAB medium (Figure 6). In addition, plants grown in the sand had thicker, shorter, and brownish roots. On the other hand, *L. japonicus* Gifu grown in a sand-vermiculite mixture nodulated with *Rl* Norway, however only when the amount of FAB medium was raised to 100 ml. No nodulation was observed with 50 ml FAB medium supplementation. These results suggest that *L. japonicus* Gifu can nodulate with *Rl* Norway conditionally, depending on the substrate moisture level or nutrient content.

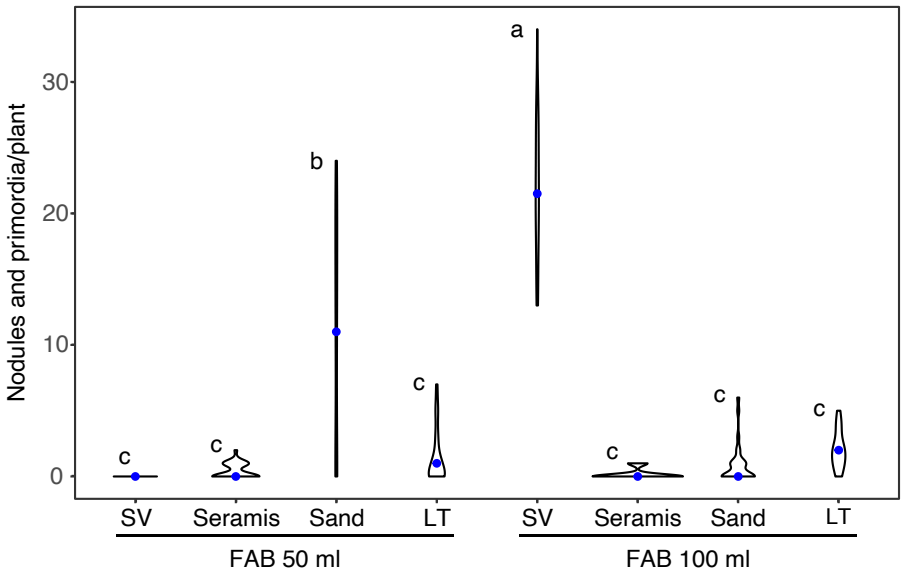


Figure 6. The nodulation of *L. japonicus* Gifu plants with *Rl* Norway under various growth conditions. The number of nodules and nodule primordia formed per plant were presented. Plants were grown in 4 distinct substrates and supplied with 50 ml or 100 ml FAB medium. Phenotyping was conducted at 6 wpi with *Rl* Norway (OD_{600} = 0.005). Thirty plants were observed for each condition. SV: sand-vermiculite; LT: Lamstedt clay granulate. Blue dots indicate the median in each condition. Different lowercase letters indicate significant differences determined by ANOVA followed by Tukey's test (p <0.05).

To specify if substrate moisture or nutrient content contributed to the nodulation of *L. japonicus* Gifu with *Rl* Norway, experiments with a gradient substrate moisture and two nutrient availabilities were performed. At 6 wpi with *Rl* Norway, the nodulation of *L. japonicus* Gifu directly correlated with the amount of FAB medium (Figure 7A). Sand-vermiculite supplemented with 100 ml FAB medium remained a nodulation condition regardless of the nutrient content, although plants grown in high substrate moisture and high nutrient content nodulated more with *Rl* Norway (Figure 7B). However, sand-vermiculite supplemented with 50 ml FAB medium remained a non-nodulation condition even when the nutrient content was raised to two times the standard quantity. These results indicate that the substrate moisture is the primary determinant for the nodulation of *L. japonicus* Gifu with *Rl* Norway. Therefore,

sand-vermiculite supplemented 50 ml FAB medium was defined as low substrate moisture (LM), whereas sand-vermiculite supplemented 100 ml FAB medium was defined as high substrate moisture (HM).

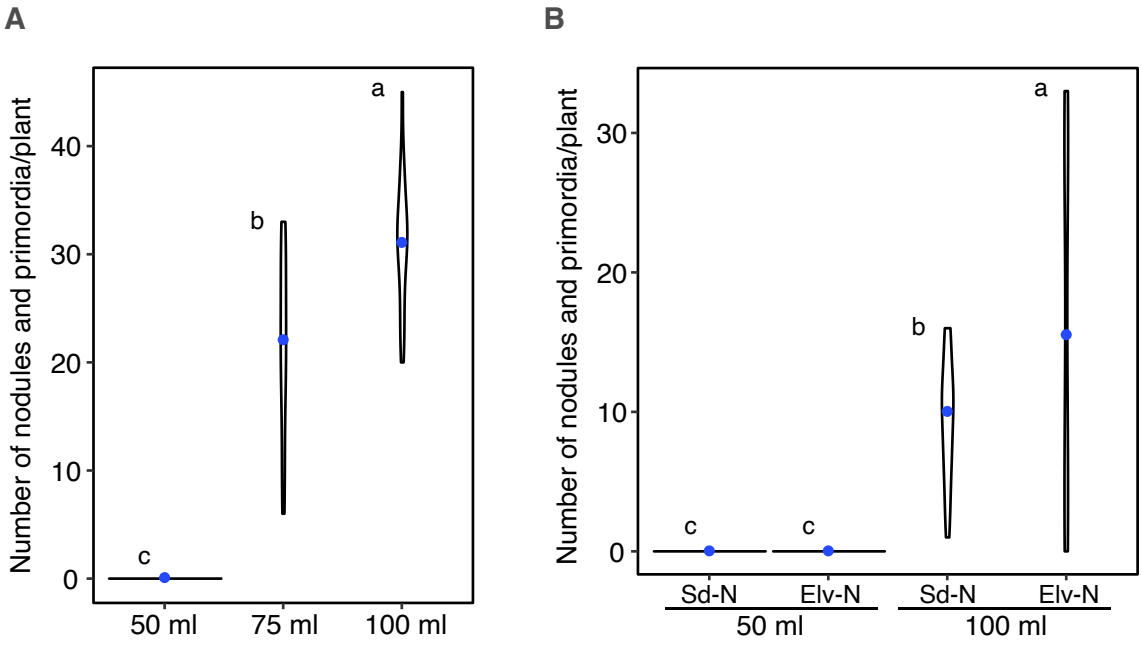


Figure 7. The effect of substrate moisture on nodulation of *L. japonicus* Gifu with *Rl* Norway. (A) Increasing amounts of FAB medium (50 ml, 75 ml, and 100 ml) were applied to plants inoculated with *Rl* Norway (OD_{600} = 0.005) in a 300 ml sand-vermiculite mixture to assess the effect of moisture. (B) Plants inoculated with *Rl* Norway (OD_{600} = 0.005) in 300 ml sand-vermiculite mixture were supplemented with 50 ml or 100 ml FAB medium in combination with a standard amount of total nutrient (Sd-N) or elevated amount to two times of total nutrient (Elv-N). Twenty individual plants were phenotyped at 6 wpi for each condition. The blue dots indicate the median of each condition. Different lowercase letters indicate significant differences determined by ANOVA followed by Tukey's test (p <0.05).

2.2 An additive effect between substrate moisture and *Rl* Norway inoculation on nodulation

In *L. japonicus* Gifu, spontaneous nodule formation has been observed in two gain-of-function mutants, *snf1-1* and *snf2-2*. The *snf1-1* and *snf2-2* mutants form nodules in the absence of rhizobia, which result from the constitutively activated CCaMK and cytokinin signaling pathways, respectively (Suzaki et al., 2013). To investigate whether the substrate moisture affects the nodule organogenesis independently of the bacteria, the nodulation of *L. japonicus* Gifu, *snf1-1*, and *snf2-2* alone or with *Rl* Norway were inspected in LM and HM conditions. At 6 wpi, *L. japonicus* Gifu, *snf1-1*, and *snf2-2* grown in LM did not show a significant difference in the number of nodules and nodule primordia between mock-treated and *Rl* Norway-inoculated plants (Figure 8). For *snf1-1* and *snf2-2*, plants grown in HM significantly developed

more nodules and nodule primordia compared with plants grown in LM. In addition, their numbers of nodules and nodule primordia were significantly increased in the presence of *RI* Norway compared to the mock groups. These results suggested an additive effect between substrate moisture and *RI* Norway inoculation.

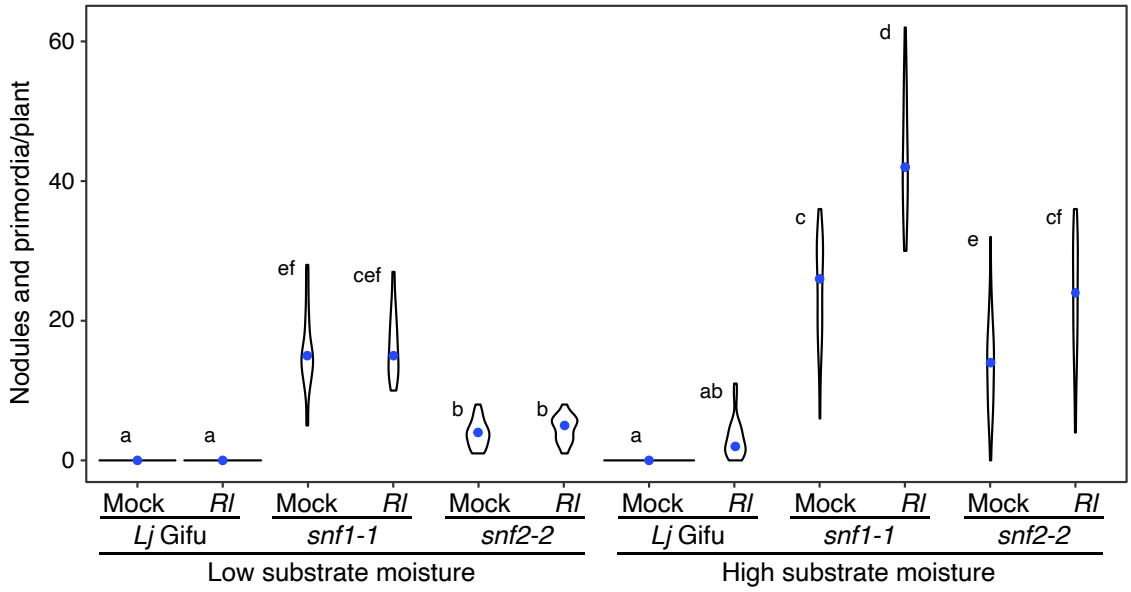


Figure 8. An additive effect between substrate moisture and *RI* Norway inoculation on nodulation of *L. japonicus* Gifu, and *snf1-1* and *snf2-2* mutants. Seedlings of wild-type *L. japonicus* Gifu (*Lj* Gifu), *snf1-1*, and *snf2-2* mutants grown in sand-vermiculite were supplemented with 50 ml FAB medium (low substrate moisture) or 100 ml FAB medium (high substrate moisture). The number of nodules and primordia on each plant was quantified at 6 wpi with *RI* Norway (OD_{600} = 0.005). For *L. japonicus* Gifu, twenty plants were assessed in each condition. For *snf1-1* and *snf2-2* mutants, thirty plants were assessed in each condition. *RI*, *RI* Norway inoculation. Different lowercase letters indicate significant differences determined by Kruskal-Wallis test followed by Dunn's post-hoc test (p <0.05).

2.3 High substrate moisture promotes the colonization of *RI* Norway

Rhizobia root colonization is one of the key contributors in the early stage of symbiosis establishment prior to nodule formation. Since *L. japonicus* Gifu nodulated with *RI* Norway in HM but not LM, the absence of nodulation in LM could be caused due to a lack of rhizobia root colonization. To investigate this, the root colonization of *RI* Norway was inspected at 2 wpi on *L. japonicus* Gifu roots grown in LM and HM. The GFP fluorescence expressed by *RI* Norway indicated that *RI* Norway presented a high proportion of surface colonization on *L. japonicus* Gifu roots grown in HM. However, *RI* Norway colonization was barely observed on *L. japonicus* Gifu roots grown in LM (Figure 9A).

To quantify the level of colonization in LM and HM, a colony-forming unit (CFU) assay and qPCR were performed. Bacteria were detached from the roots by incubating in sterile water for at least 16 h prior to conducting CFU assay and qPCR. The CFU quantification indicated the cell number of bacteria per root fragment in HM and LM was 2.4e+06 and 5.96e+02 on average, respectively (Figure 9B). A qPCR was performed with the primers targeting the 16S rRNA of *Rl* Norway (Table S4). The *Rl* Norway cell number per root was calculated by the calibration curve generated by a serial dilution of *Rl* Norway genomic DNA (Figure 9C). The qPCR quantification showed the average number of *Rl* Norway cells per root fragment was 9.7e+07 in HM and 9.7e+01 in LM (Figure 9D). Both CFU and qPCR quantifications confirmed that *L. japonicus* Gifu grown in HM supported higher *Rl* Norway root colonization. Hence, HM promoted the root colonization by *Rl* Norway.

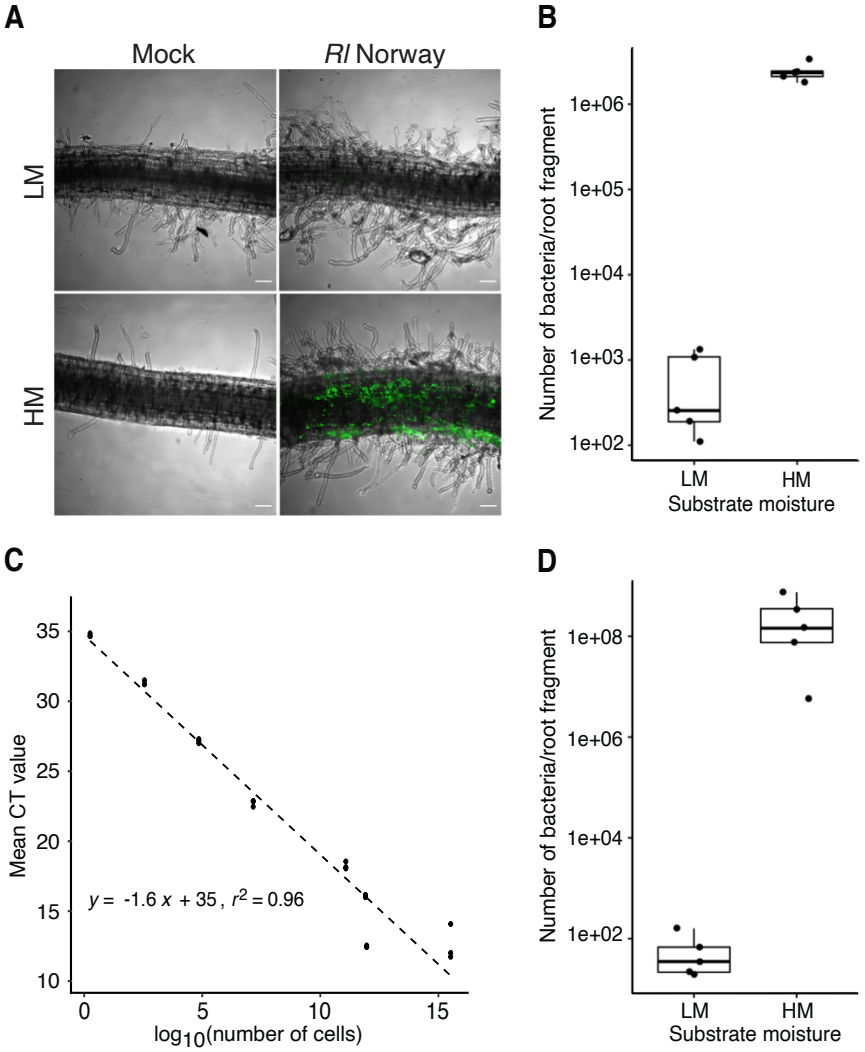


Figure 9. *RI* Norway presented higher root colonization in high substrate moisture. Plants were grown in sand-vermiculite substrate supplemented with 50 or 100 ml of FAB medium and examined at 2 wpi with *RI* Norway. (A) Representative images of root colonization of *L. japonicus* Gifu under low and high substrate moisture conditions. At least three plants from three pots were analyzed for each treatment in two independent experiments. Scale bar = 100 μm. (B) Quantification of *RI* Norway colonization by colony-forming unit. Root fragments, 3 cm from the tip of the primary roots, were collected from five different pots. *RI* Norway was detached from *L. japonicus* Gifu before quantification by incubating the roots in sterile water overnight, sonication at 50 kHz for 15 min, followed by vortexing for 2 min. (C) The calibration curve for *RI* Norway quantification by qPCR. The standard curve was generated using a ten-fold serial dilution of *RI* Norway gDNA. For each dilution, three technical replicates of qPCR reactions were performed. The number of cells was calculated based on the molecular weight of *RI* Norway (7,788.085 kb in genome size) (Liang et al., 2018). As 1 kb of DNA weighs approximately 10^{-6} pg (Sessions, 2013), each *RI* Noway cell contains approximately 7.79 x 10^{-6} ng of DNA. Thus, 1 ng of gDNA is equivalent to approximately 1.28 x 10^{5} cells of *RI* Norway. (D) Quantification of *RI* Norway colonization by qPCR. Bacterial DNA was extracted from the root fragments collected from five different pots. qPCR was performed with *RI* Norway 16S rRNA gene primers (Table S4). Each dot represents the mean of three technical replicates from one biological replicate.

3 Using RNA-seq to investigate the difference in the regulation of symbiotic compatibility between *L. burttii* and *L. japonicus* Gifu

A greater variation in symbiotic compatibility between sub-compatible partnerships was observed, where *L. burttii* and *L. japonicus* Gifu showed the most and the least restrictive symbiont selection. Additionally, a moister growth environment enhanced the symbiotic compatibility between *L. japonicus* Gifu and *Rl* Norway and led to nodulation. These findings raised two biological questions: which plant genes determine symbiotic compatibility and what are the mechanisms responsible for the compatibility alteration by HM.

To determine the difference in the transcriptomic response to *Rl* Norway in the different hosts, we adapted the molecular crowding single-cell RNA barcoding and sequencing (mcSCRB-seq) method (Bagnoli et al., 2018). The most and the least restrictive *Lotus* accessions, *L. burttii* and *L. japonicus* Gifu were grown in LM. Plants were either inoculated with *Rl* Norway or treated with the FAB medium as a mock control. To uncover how high moisture alters the symbiotic compatibility in *L. japonicus* Gifu, additional *L. japonicus* Gifu plants were grown in HM and either inoculated with *Rl* Norway or treated with the FAB medium. Six biological replicates were prepared for each condition, and three additional biological replicates were prepared for mock groups of *L. japonicus* Gifu grown in LM and HM. Each biological replicate contained roots from three individual plants. Root RNA was reverse-transcribed to cDNA with sample-specific bar codes and labeled with unique molecular identifiers (UMIs) for identifying unique molecules and minimizing the noise produced in the cDNA amplification (Ziegenhain et al., 2018). After 3'-end sequencing, the raw data was filtered with stringent criteria using the zUMI pipeline (Parekh et al., 2018) and mapped to the reference genome of *L. japonicus* Gifu v1.2 (Kamal et al., 2020) with splice-aware aligner STAR (Dobin et al., 2013).

Total transcriptomic reads were classified into mapping categories such as exon, intron, intergenic, ambiguity, unmapped and unused barcode (Figure 10A). More than 75% of the reads were mapped to the *L. japonicus* Gifu v1.2 genome (Kamal et al., 2020), while less than 10% were ambiguous (Figure 10B). The reads were categorized into 1,324,152 UMIs. They were mapped to the genome and assorted into three categories: exon, intron, and the overlap between the exon and intron. A number of 646,296 UMIs represented 22,774 exons of genes, 17,926 UMIs represented 3,281 introns of genes, and 659,930 UMIs represented 23,150 overlaps between the exon and intron (Figure 10C). The sum of the detected genes in the exon and intron represented approximately 86% of 30,243 annotated genes in the *L. japonicus* Gifu v1.2 genome (Kamal et al., 2020). In each biological replicate, reads were mostly mapped to exon, intergenic region, and intron. The pattern of reads distribution was consistent in all

samples (Figure 10D). Altogether, the expression of the majority of genes in the genome of *L. japonicus* Gifu was detected by RNA-seq.

To identify the similarity between each sample, principal component analysis (PCA) was performed. Samples of *L. burttii* and *L. japonicus* Gifu inoculated with *Rl* Norway clustered into three different groups by principal components 1 and 2 (PC1 and PC2), which contributed 38% and 30 % of the variance. The mock samples of *L. burttii* and *L. japonicus* Gifu grown in HM also clustered together, whereas the mock samples of *L. japonicus* Gifu grown in LM were relatively dispersed. Excluding one mock sample from *L. japonicus* Gifu grown in HM and one mock sample from *L. japonicus* Gifu in LM due to low total read counts, all samples were included for further analysis.

To identify what genes are regulated in response to *Rl* Norway, the DEGs between mock and *Rl* Norway inoculated samples were determined using the DEseq2 package (Love et al., 2014) in R and illustrated in a Venn Diagram (Figure 11B, Supplementary File 1). The number of DEGs between mock and *Rl* Norway inoculation was 465 in *L. burttii*, 763 in *L. japonicus* Gifu grown in LM, and 1,096 in *L. japonicus* Gifu grown in HM. Among the DEGs, 426, 121, and 436 genes were exclusively present in *L. burttii*, *L. japonicus* Gifu grown in LM, and *L. japonicus* Gifu grown in HM, respectively. Genes shared between two comparisons were 657 in sum. Among those, 627 genes were shared between *L. japonicus* Gifu grown in LM and HM. Moreover, 9 genes were shared between the three comparisons (Figure 11B). Further computing the distance between each sample showed a species-wise difference between *L. burttii* and *L. japonicus* Gifu (Figure 11C). The hierarchical clustering demonstrated that samples under the same condition were grouped and the color similarity was species-distinguishable (Figure 11C). As RNA-seq data presented the differences between plant species and conditions, these data were used for further analyses.

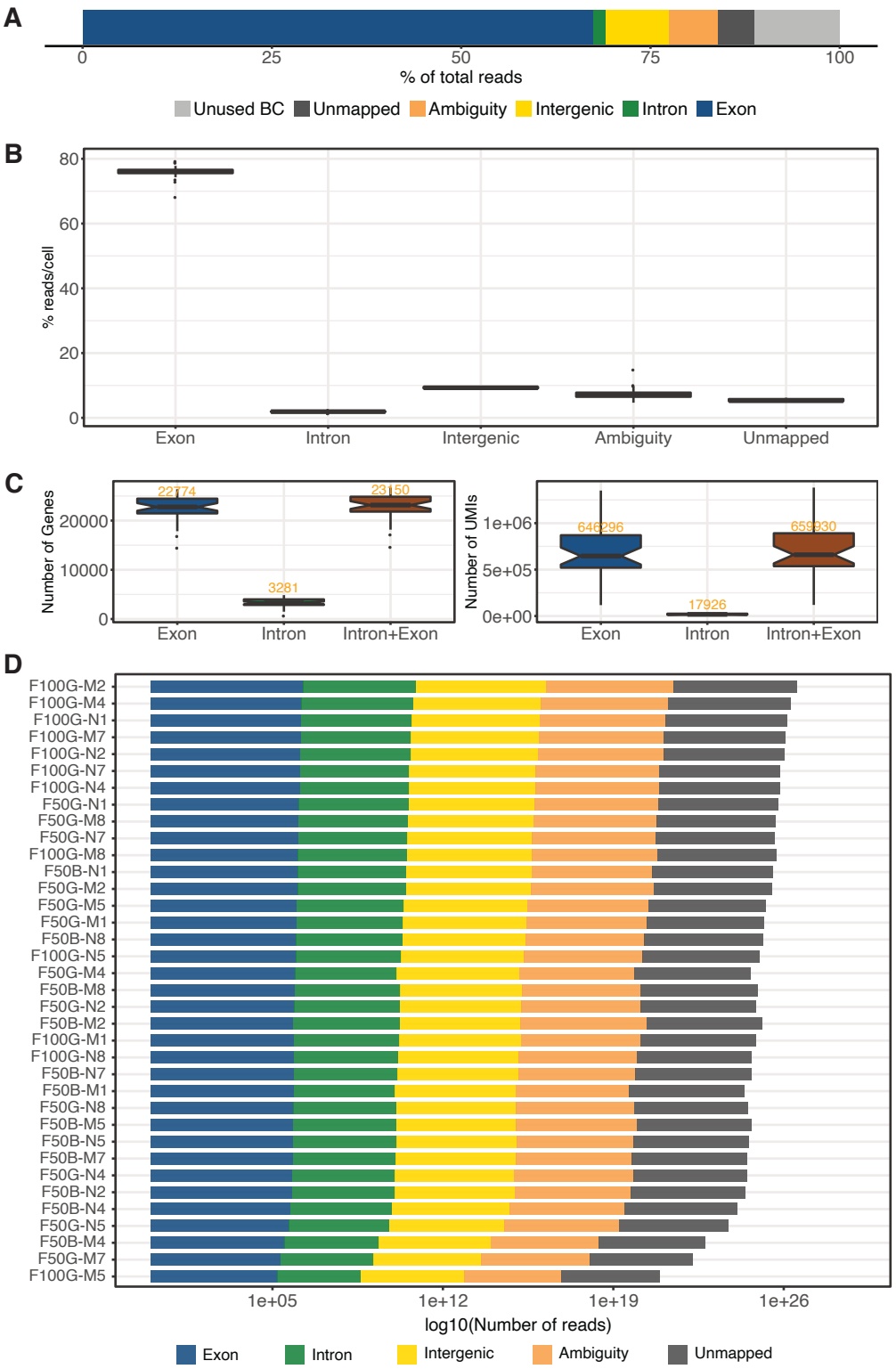


Figure 10. Quality measurements of the transcriptome sequencing. Transcriptome sequencing (RNA-seq) was conducted with *L. japonicus* Gifu (G) and *L. burttii* (B) roots at 2 wpi with mock (M) or *Rl* Norway (N) under low (F50) and high (F100) substrate moisture conditions. (A) and (B) Distribution of the total reads per barcoding cell in different genomic regions as well as the ambiguous and unmapped reads in percentage. BC: barcode. (C) The distribution of the detected genes and UMIs in the exon, intron, and the overlapped region between the intron and exon. (D) Distribution of reads for each biological replicate of the six conditions. The data was generated using the zUMIs pipeline (2.5.4) (Parekh et al., 2018) with STAR (2.6.0) (Dobin et al., 2013).

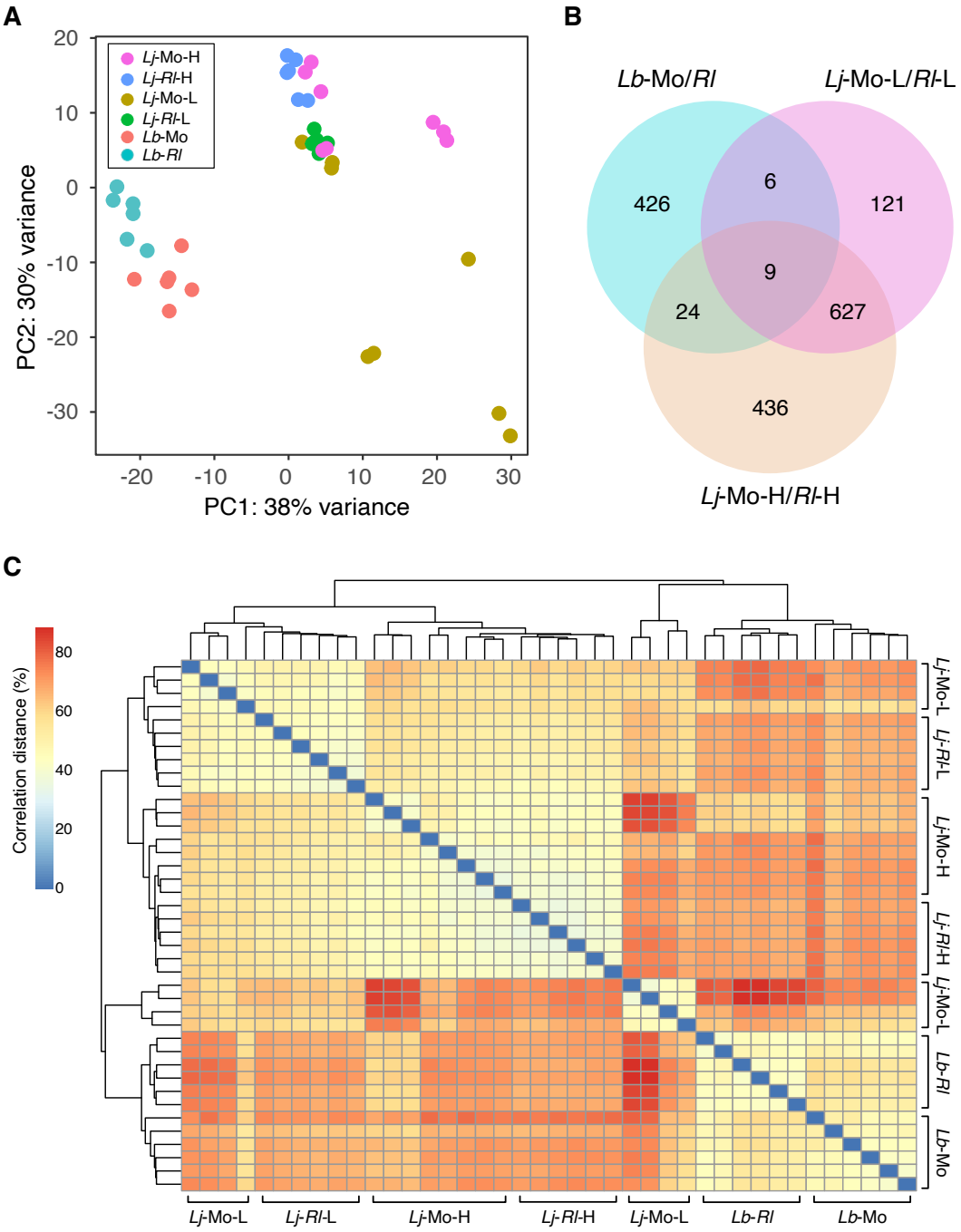


Figure 11. Global analysis of transcriptomic sequencing. (A) Principal component analysis (PCA) of *L. burttii* (*Lb*) and *L. japonicus* Gifu (*Lj*) with mock treatment (Mo) and *R. leguminosarum* Norway (*Rl*) inoculation in combination with low (L) and high (H) substrate moisture. PC1: principal component 1, PC2: principal component 2. (B) Numbers of genes differentially expressed in comparisons within plant species and growth conditions (Mo versus *Rl*). The Venn diagram shows the intersection of genes differentially expressed in each comparison (p <0.01). (C) Sample correlation. Biological replicates of each examined condition were analyzed and categorized by sample similarity, which is represented by. correlation distance. The correlation distance is 0 in the diagonal position as identical samples are compared. The analyses were conducted in R with the DEseq2 package (Love et al., 2014).

4 Investigating the regulatory mechanism underlying the nodulation of *L. japonicus* Gifu in high substrate moisture

4.1 Genes regulated by substrate moisture have a putative function in the organonitrogen compound metabolism and oxidation-reduction processes

To explore how substrate moisture contributes to the nodulation of *L. japonicus* Gifu with *RI* Norway at the transcriptomic level, transcriptomic profiles of *L. japonicus* Gifu grown in LM and HM under mock treatment and *RI* Norway inoculation conditions were analyzed. DEGs were determined by the DESeq2 package (Love et al., 2014) in R using a p <0.01. The expression of 3,101 and 1,799 genes was altered by substrate moisture in mock-treated and *RI* Norway-inoculated roots, respectively (Figure 12A, Supplementary File 2). Among them, 2,708 genes were differentially expressed only in mock roots (Figure 12A, G1), 1,406 genes were differentially expressed exclusively in *RI* Norway inoculated roots (Figure 12A, G3), and 393 genes were differentially expressed in both treatments (Figure 12A, G2). Within these DEGs, the expression of 1,562 genes was changing over two-fold (absolute value of $\log_2$ fold change >1), including 1,018 genes from mock-treated roots (Figure 12A, G1), 527 genes from *RI* Norway-inoculated roots (Figure 12A, G3), and 17 genes were shared in both treatments (Figure 12A, G2). Overall, these DEGs were classified as genes regulated by substrate moisture (Figure 12A and Supplementary File 3, G1+G2) and genes that were specifically regulated by *RI* Norway in addition to substrate moisture (Figure 12A and Supplementary File 3, G3).

To clarify the DEGs into functional categories, gene ontology (GO) analyses based on biological processes were performed on the DEG commonly regulated by substrate moisture (Figure 12A and Supplementary File 3, G1+G2, p <0.01) and the DEGs specifically regulated by *RI* Norway in addition to substrate moisture (Figure 12A and Supplementary File 3, G3, p <0.01), respectively. The organonitrogen compound metabolic process (GO:1901564) was the most over-represented GO category for genes commonly regulated by substrate moisture (Figure 12B), followed by the cellular protein metabolic process (GO:0044267), gene expression (GO:0010467), and organonitrogen compound biosynthetic process (GO:19015646). Several GO categories related to organonitrogen compound metabolism and biosynthesis were also present, including amide and peptide biosynthesis processes (GO:0043604 and 0043043) as well as amide and peptide metabolism processes (GO:0043603 and 0006518). The GO category – translation (GO:0006412) was a part of gene expression (GO:0010467). In these GO categories, genes encoding ribosomal proteins, elongation factors, translation initiation factors, and protein kinases were the most representative genes. For the DEG in group 1 and group 2 (Figure 12A, G1+G2) with p <0.01

and absolute value of $\log_2$ fold change >1, the peptide metabolic process (GO:0006518) was the most over-representative GO category. Altogether, genes commonly involved in cellular organonitrogen compound metabolic processes were differentially expressed between LM and HM.

For DEG specifically affected by *RI* Norway inoculation (Figure 12A, G3, p <0.01), the oxidation-reduction process (GO:0055114), transport (GO: 0006810), and membrane transport (GO:0055085) were the most over-represented GO categories (Figure 12C). Although the annotated gene number was relatively low, other GO categories included the response to oxidative stress (GO: GO:0006979), the reactive oxygen species metabolic process (GO:0072593), and its subcategories – hydrogen peroxide catabolism (GO: 0042744) and hydrogen peroxide metabolism (GO:0042743). Three cellular metabolic processes-related GO categories were represented as well, comprising dephosphorylation (GO:0016311), antibiotic catabolic process (GO:0017001), and cofactor catabolic process (GO:0051187). Genes encoding peroxidases, transporters, cytochrome P450 family proteins, and enzymes were the majority in these GO categories. In addition, the oxidation-reduction process (GO:0055114) was also the most over-represented GO category of the DEG in group 3 with $p < 0.01$ and absolute $\log_2$ fold change >1 (Figure 12A, G3). In summary, the expression of genes functionally associated with oxidation-reduction processes was altered after *RI* Norway inoculation.

4.2 Co-expressed genes regulated by the substrate moisture after *RI* Norway inoculation were also over-represented in the oxidation-reduction process

To investigate the relationship between DEGs and narrow down the candidates, a Weighted Gene Co-Expression Network Analysis (WGCNA) was conducted. Transcriptome datasets available from the Lotus base (https://lotus.au.dk/) were extracted from plant organs (mature flower, immature flower, pod, seed, leaf, and root), from roots treated with either pathogenic (*Ralstonia solanacearum* JS763, *Pseudomonas syringae* pv. tomato DC3000) or symbiotic (*Ef* HH103, *Bradyrhizobium elkanii* USDA61, and *M. loti* R7A) bacteria and nodules generated from plants inoculated with *M. loti* R7A at 7 and 10 days post-inoculation (dpi). Five co-expression modules and one module comprising uncategorized genes were produced (Figure 12D, Supplementary File 4), which included 401, 326, 323, 223, 113, and 18 genes.

GO analyses performed for genes in each module showed that module 1 was over-represented by cellular metabolic processes (Figure 13A), including the phosphate-containing compound metabolic process (GO:0006796), phosphorus metabolic process (GO:0006793), cell

communication (GO:0007154), and carbohydrate biosynthetic process (GO:0016051). The trehalose biosynthetic process (GO:0005992) was the representative child category of the carbohydrate biosynthetic process. The response to ethylene category (GO:0009723) was shown in the result, which is a child term of the response to stimulus (GO:0050896). In module 2, the most over-represented GO term was the oxidation-reduction process (GO:0055114). Its closely related categories could also be found in the result, including the reactive oxygen species metabolic process (GO:0072593) and hydrogen peroxide metabolic and catabolic process (GO:0042743 and 0042744). In addition, the GO analysis result included four drug- and antibiotic-related catabolic and metabolic processes (GO:0042737, 0017144, 0017001, and 0016999). Genes in module 3 were over-represented by the regulation of biological process (GO:0050789) and its child term, the regulation of cellular process (GO:0050794), followed by the transport (GO:0006810) and its child term, the transmembrane transport (GO:0055085). Other minor categories for the genes in module 3 contained the photosynthesis (GO:0015979), plant-type cell wall organization or biosynthesis (GO:0071669), protein repair (GO:0030091), alcohol biosynthetic process (GO:0046165), and ammonium-related transportation categories (GO:0015696 and 0072488). Similar to module 2, the oxidation-reduction process remained the most over-represented category in module 4 (Figure 13D). Only three genes were assigned to the cellular ion homeostasis (GO:0006873), cation homeostasis (GO:0055080), inorganic ion homeostasis (GO:0098771), and their subcategories. For genes in module 5, the carbohydrate metabolic process (GO:0005975) was the most over-represented category, followed by the alpha-amino acid metabolic process (GO:1901605). The monosaccharide metabolic process (GO:0005996) was the child term of the carbohydrate metabolic process, while the glutamine family amino acid metabolic process (GO:0009064) and proline catabolic process (GO:0006562) were subcategories of the alpha-amino acid metabolic process. In addition, module 5 consisted of the organ growth-related categories (GO:0035265, 0045926, 0046620, and GO:0046621). However, only few genes were categorized in each of the GO categories (Figure 13E).

Overall, the oxidation-reduction process was the most over-represented and significant GO category of the six co-expressed gene modules (Figure 13). The oxidation-reduction process was also the most significant term of the DEG that was specifically regulated by *RI* Norway in addition to substrate moisture (Figure 12B). It is not surprising that oxidation-reduction reactions stood out in the analysis as they are involved in numerous biochemical processes regulating plant metabolism. A higher number of phosphate containing compound metabolism, carbohydrate metabolism, and secondary metabolic processes suggested a potential role of secondary metabolites.

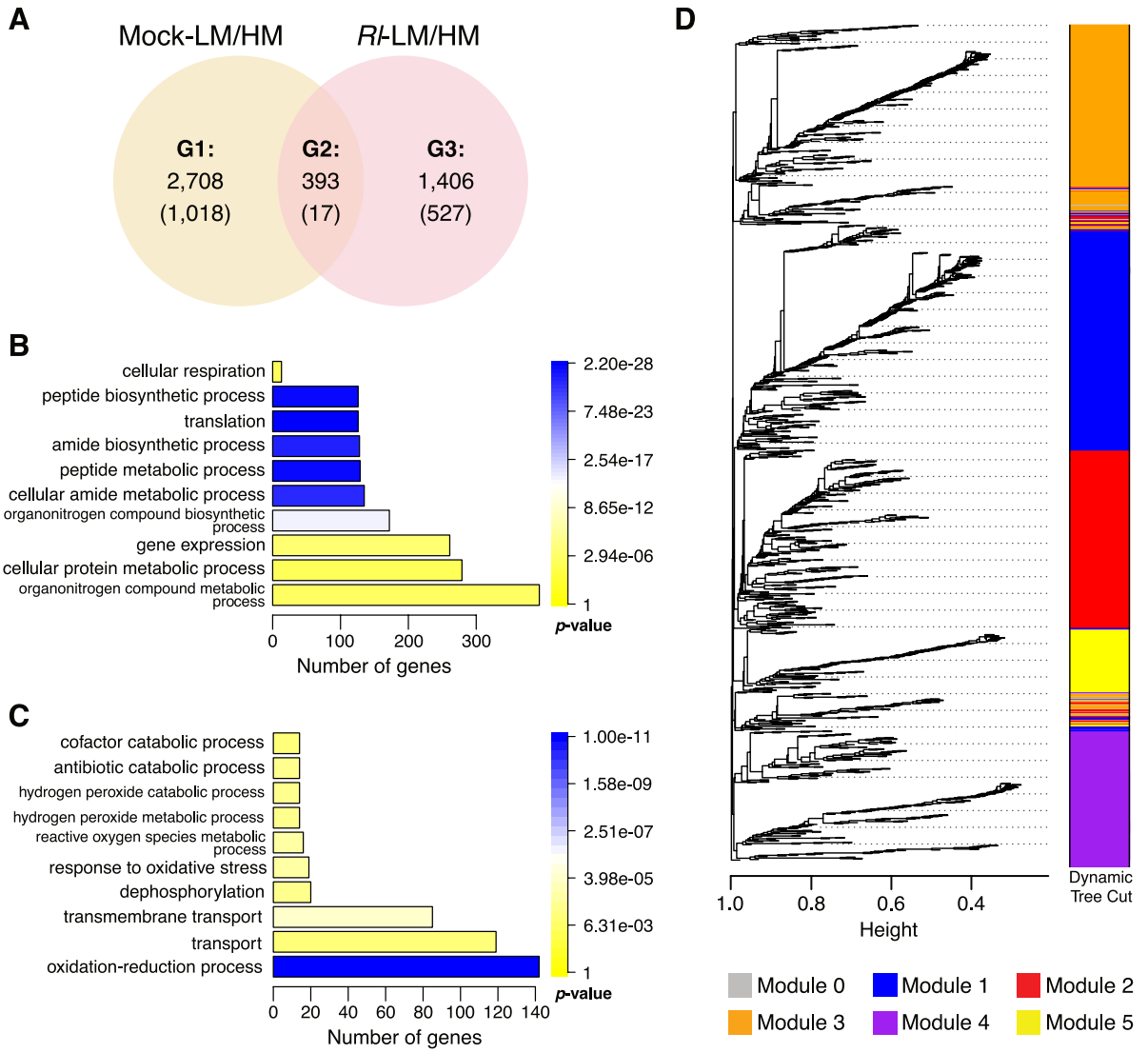


Figure 12. Analysis of differentially expressed genes in *L. japonicus* Gifu between low and high substrate moisture under mock treatment and *RI* Norway inoculation conditions. (A) Numbers of genes regulated by substrate moisture after mock treatment and *RI* Norway inoculation. Differentially expressed genes (DEGs) between *L. japonicus* Gifu grown in low and high substrate moisture were determined by the DEseq2 package in R (Love et al., 2014). The upper numbers indicate the number of DEG with a p <0.01. The lower numbers in the brackets indicate the number of genes with p <0.01 and absolute value of $\log_2$ fold change >1. LM, low substrate moisture; HM, high substrate moisture; *RI*, *RI* Norway inoculation; G, groups of genes in corresponding comparisons. (B) and (C) Gene ontology (GO) categories of DEG (p <0.01) in G1 and G2 (B), and G3 (C), respectively. The analysis was performed in R with the topGO package (Alexa & Rahnenfuhrer, 2019). The DEGs in each GO category and their functional annotations are listed in Supplementary File 3. (D) Dendrogram of the co-expressing modules generated by genes with G3 (p <0.01). Different colors depict genes in different co-expression groups. The transcriptomic profiles were acquired from Expression Atlas of Lotus Base (https://lotus.au.dk/expat). A weighted correlation network analysis was conducted with the WGCNA package in R (Langfelder & Horvath, 2008).

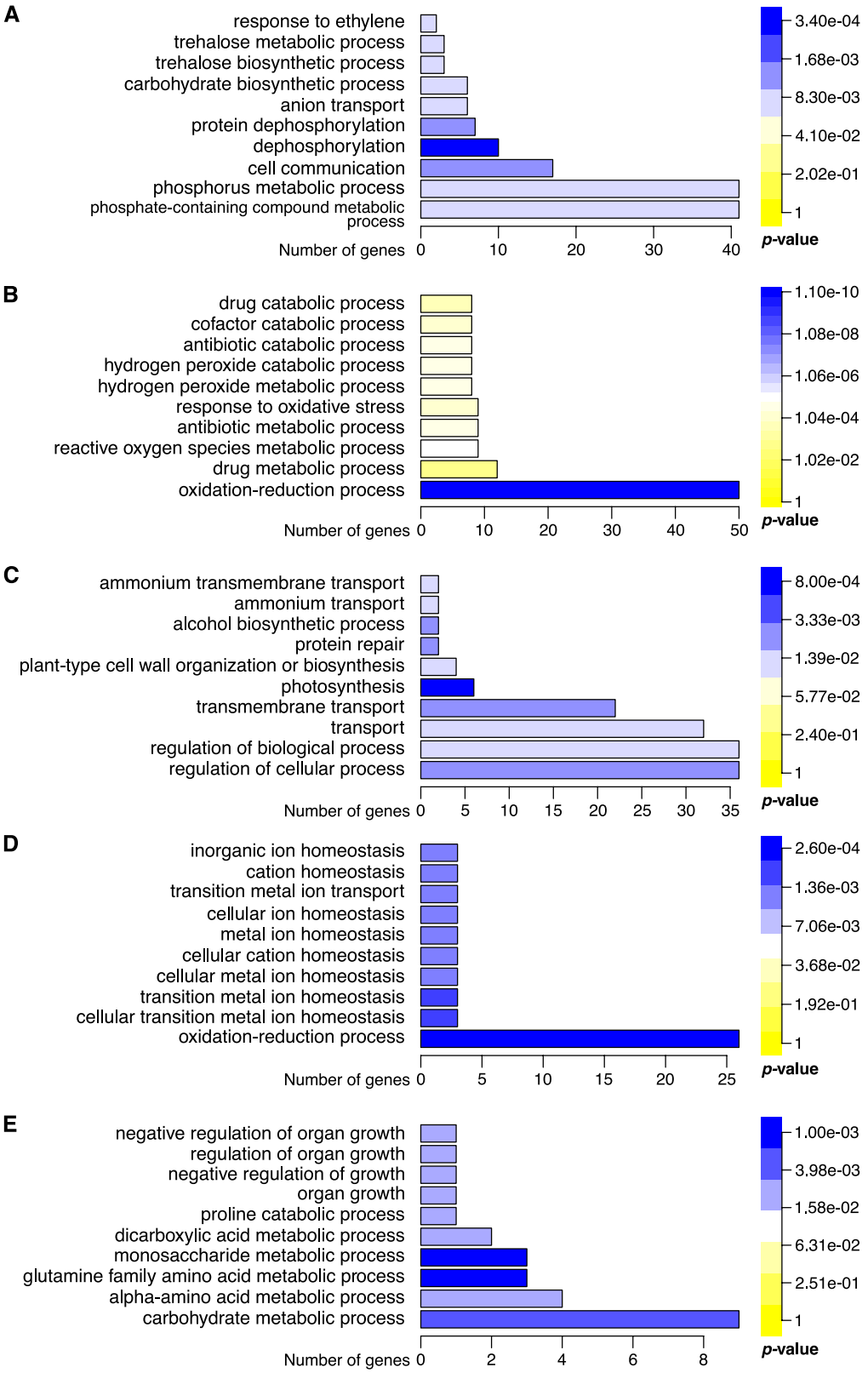


Figure 13. Gene ontology analysis of co-expressed gene modules in *L. japonicus* Gifu between low and high substrate moisture after *RI* Norway inoculation. Gene ontology (GO) analysis was applied to the differentially expressed genes (DEGs) in each co-expression module using the topGO package in R (Alexa & Rahnenfuhrer, 2019). (A) to (E) Over-represented GO categories in modules 1 to 5 of DEGs specifically regulated by substrate moisture after *RI* Norway inoculation, respectively.

4.3 Phytohormone-related genes were present in the GO category of the oxidation-reduction process

As the oxidation-reduction process (GO:0055114) was the most over-represented GO category not only for genes regulated by substrate moisture after *Rl* Norway inoculation but also for their co-expressed gene clusters, we further explored the genes in module 2 and module 4 (Figure 13B, D and Supplementary File 4) that were grouped in the oxidation-reduction process category. Most of the genes in this category of module 2 encoded peroxidases, cytochrome P450 family proteins, and oxygenases. Additionally, genes encoding 1-aminocyclopropane-1-carboxylate oxidase (ACO), 12-oxophytodienoate reductase (OPR), and cytokinin dehydrogenase 3 (CKX3) were found. These genes were involved in ethylene biosynthesis (Houben & Van de Poel, 2019), jasmonic acid biosynthesis (Zdyb et al., 2011), and cytokinin degradation (Reid et al., 2016), respectively. Similarly, genes in the oxidation-reduction process category of module 4 encoded cytochrome P450 family proteins, oxygenases, and dehydrogenases. In addition, a gibberellin 3-beta-hydroxylase (GA3OX2) encoding gene was identified, which participates in GA biosynthesis and converts GA precursors into bioactive GAs (Yamaguchi et al., 1998). The transcriptomic data showed that *ACO* and *OPR* were downregulated in HM, and *GA3OX2* and *CKX3* were upregulated in HM (Figure 14). Although these results do not demonstrate the direct involvement of phytohormones in the nodulation in HM, they suggest that phytohormone metabolisms may be fine-tuned by substrate moisture after *Rl* Norway inoculation.

As the ACO is the key and rate-limiting enzyme that catalyzes the last step in ethylene biosynthesis (Houben & Van de Poel, 2019) and numerous studies have demonstrated the role of ethylene in symbiosis, especially its negative role in regulating nodulation since the 1970s (Guinel, 2015, Grobbelaar et al., 1971), we further investigated the effect of ethylene on the nodulation of *Rl* Norway. To assess if ethylene contributes to nodulation either positively or negatively, 1 µM 1-aminocyclopropane-1-carboxylic acid (ACC, ethylene precursor) and 1 µM aminoethoxyvinylglycine (AVG, ethylene biosynthesis inhibitor) were applied to *L. japonicus* Gifu seedlings grown in LM and HM and inoculated *Rl* Norway. At 6 wpi, *L. japonicus* Gifu still nodulated under HM conditions, and no significant difference in the number of organogenesis events was detected (Figure 15).

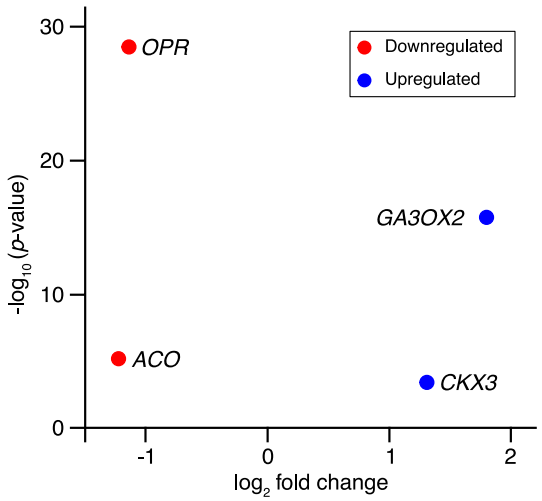


Figure 14. Genes involved in phytohormone synthesis and catabolism were regulated by substrate moisture after *RI* Norway inoculation. The root transcriptomes of *L. japonicus* Gifu grown in LM and HM for 14 days after *RI* Norway inoculation were inspected. Fold change in the genes is relative to the LM condition. The differentially expressed genes involved in plant hormone biosynthesis and catabolism that were classified in the oxidation-reduction process category by gene ontology analysis were presented. Upregulated and downregulated genes are depicted in blue and red, respectively. LM, low substrate moisture; HM, high substrate moisture; *ACO*, *1-aminocyclopropane-1-carboxylate oxidase*; *OPR*, *12-oxophytodienoate reductase*; *CKX3*, *cytokinin oxidase/dehydrogenase 3*; *GA3OX2*, *gibberellin 3-beta-hydroxylase 2*.

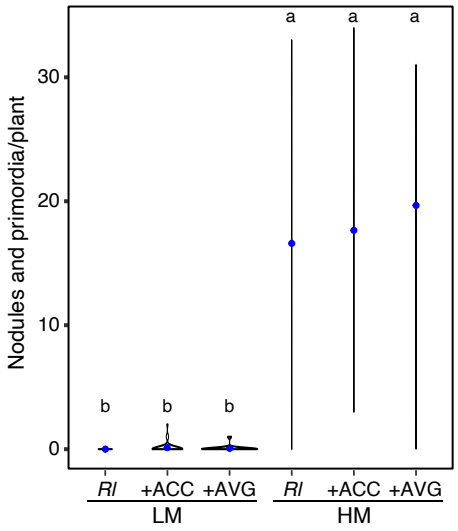


Figure 15. The effect of ethylene precursor and inhibitor on the nodulation of *L. japonicus* Gifu with *RI* Norway in low and high substrate moisture. *L. japonicus* Gifu seedlings grown in low and high substrate moisture were inoculated with *RI* Norway (*RI*, OD_{600} = 0.005) and supplied with additional 0.5 µM 1-aminocyclopropane-1-carboxylic acid (ACC, ethylene precursor) or 1 µM aminoethoxyvinylglycine (AVG, ethylene biosynthesis inhibitor). Twenty plants were inspected for conditions without ACC and AVG application. Thirty plants were inspected for each condition with ACC or AVG application. Different lowercase letters indicate significant differences determined by ANOVA followed by Tukey's test (p <0.05). LM, low substrate moisture; HM, high substrate moisture.

4.4 Genes involved in flavonoid biosynthesis were downregulated in high substrate moisture condition

Genes involved in secondary metabolites synthesis were also observed in the oxidation-reduction process category (GO:0055114) of module 2 and module 4 (Figure 13B and D). In addition to the genes encoding 2-oxoglutarate and Fe(II)-dependent oxygenases, several genes involved in flavonoid biosynthesis were identified. The flavonoid biosynthesis pathway is derived from the phenylpropanoid biosynthesis pathway, which also derives into coumarin biosynthesis (Figure 16A). Flavonoids can induce the transcription of Nod factor biosynthesis genes, function as antioxidants against abiotic stress, and are required in plant defense against pathogen and herbivores (Dong & Lin, 2021). Genes encoding chalcone synthase (CHS) mediate the first step of flavonoid and isoflavonoids biosynthesis. Together with chalcone isomerase (CHI), they generate an intermediate, naringenin, from *p*-coumaric acid. Flavanone 3-hydroxylase (F3H) and flavonol synthase (FLS) then transform naringenin into flavonoid end products (Figure 16A). Among them, *CHS*, *F3H*, and *FLS* were downregulated over two-fold in HM (Figure 16B). Additionally, a gene encoding polyketide reductase (PKR), which functions at the second step of isoflavonoids biosynthesis, was also downregulated in HM (Figure 16B). As the phenylpropanoid, coumarin, flavonoid and isoflavonoids biosynthesis pathways are connected and many enzymes in those biosynthesis pathways are encoded by a group of genes, we further investigated the expression level of such gene groups. These gene groups were identified by searching in the protein databases with respective protein queries. For instance, we found a group of 9 genes that encode phenylalanine ammonia lyase (PAL) and a group of 3 genes that can encode FLS in *L. japonicus* Gifu. Two genes downregulated in HM were identified as a putative feruloyl-CoA 6′-Hydroxylase 1 (F6’H1)- and coumarin synthase (COSY)-encoding gene after protein blast, respectively. However, a total of 48 inspected genes were not differentially regulated between LM and HM. Overall, three out of four key enzymes were downregulated in HM (Figure 16B), suggesting a role of flavonoids in high moisture-dependent nodulation of *L. japonicus* Gifu with *Rl* Norway. The inspected genes were listed in Table S8.

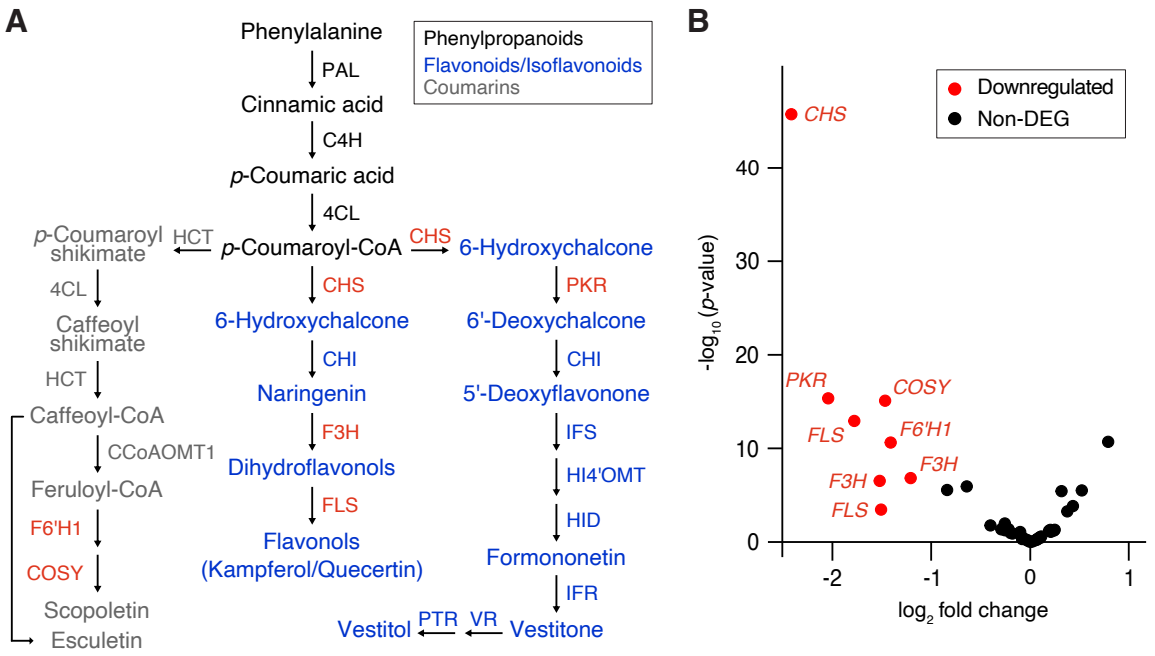


Figure 16. Differentially expressed genes in the phenylpropanoid biosynthesis pathway. (A) Phenylpropanoid biosynthesis pathway (Dong & Lin, 2021, Garcia-Calderon et al., 2020). Enzymes and intermediates in the general phenylpropanoid biosynthesis pathway are depicted in black. Enzymes and intermediates in flavonoid and isoflavonoid biosynthesis pathways are depicted in blue. Enzymes and intermediates in the coumarin biosynthesis pathway are depicted in grey. (B) The relative expression level of genes involved in flavonoids and isoflavonoids biosynthesis. RNA-seq data were obtained from roots of *L. japonicus* Gifu grown in low and high substrate moisture for 14 days after *Rl* Norway inoculation. Fold change in the genes was relative to low substrate moisture condition. Genes downregulated in high substrate moisture are depicted in red and non-differentially expressed genes (non-DEG) are depicted in black. The DEG analysis was conducted by R with the DESeq2 package (Love et al., 2014). PAL, phenylalanine ammonia lyase; C4H, cinnamic acid 4-hydroxylase; 4CL, 4-cummarate-CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3 - hydroxylase; FLS, flavonol synthase; PKR, polyketide reductase; IFS, 2-hydroxyisoflavonone synthase; HI4'OMT, 2-hydroxyisoflavonone 4'-O-methyltransferase, HID, 2-hydroxyisoflavonone dehydratase; IFR, isoflavone reductase; PTR, pterocarpan reductase; VR, vestitone reductase; HCT, hydroxycinnamol-coenzyme A shikimate; 4CL, 4-coumarate-CoA ligase; F6'H1, feruloyl-CoA 6′-Hydroxylase 1; COSY, coumarin synthase.

4.5 Naringenin promotes *Rl* Norway colonization but not nodulation of *L. japonicus* Gifu by *Rl* Norway

The transcriptome showed the genes encoding CHS, F3H, and FLS in the flavonoid biosynthesis pathway were repressed in HM (Figure 16), which could lead to either the accumulation of *p*-coumaroyal-CoA, naringenin, or the reduction of flavonols. While the function of *p*-coumaroyal-CoA in symbiosis remains obscure, naringenin has shown the ability to induce the *nodA* promoter activity of *R. leguminosarum* bv. *viciae* (Zaat et al., 1987). We hypothesized that naringenin accumulated in HM, and this accumulation of naringenin may

influence the bioactivity of *Rl* Norway, increase the transcription of *nod* genes in *Rl* Norway and further promote the nodulation of *L. japonicus* Gifu with *Rl* Norway in HM.

To inspect whether naringenin affects the growth of *Rl* Norway, a growth curve of *Rl* Norway with three different concentrations of naringenin was measured. No growth inhibition was detected at 0.1 µM and 1 µM of naringenin, but 10 µM naringenin reduced the growth of *Rl* Norway after 10 hours of incubation (Figure 17A). Hence, the inhibitory effect of naringenin on the growth of *Rl* Norway was concentration-dependent.

To inspect whether naringenin has an effect on the nodulation of *L. japonicus* Gifu with *Rl* Norway, *L. japonicus* Gifu plants inoculated with *Rl* Norway were grown in LM and HM either in the absence or in the presence of 1 µM naringenin. At 6 wpi, plants grown in LM showed no nodulation, whereas plants grown in HM nodulated. The difference in numbers of nodules and nodule primordia was not significant between groups with and without naringenin addition (Figure 17B). The findings indicated that naringenin did not promote nodulation under the tested conditions.

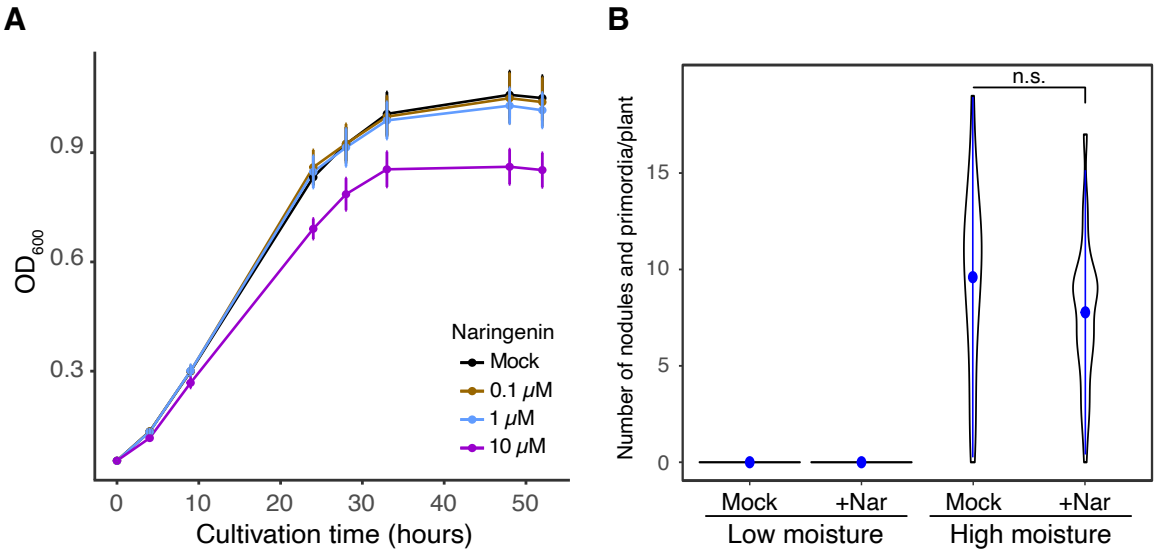


Figure 17. The effect of naringenin on the growth of *R. leguminosarum* Norway and nodulation of *L. japonicus* Gifu. (A) The growth curve of *Rl* Norway. *Rl* Norway was cultivated in 50 ml TY liquid media supplemented with streptomycin 500 µg/ml, tetracyclin 2 µg/ml, and 0.1, 1, or 10 µM of naringenin at 28ºC, 180 rpm for 3 days. The *Rl* Norway suspension had an OD_{600} = 0.054 at time 0. Bars represent the standard deviation of 3 measurements from 3 technical replicates of each treatment. (B) Nodulation of *L. japonicus* Gifu with *Rl* Norway (OD_{600} = 0.005) supplemented with 1 µM naringenin (+Nar) at 6 wpi. Plants were grown in low and high substrate moisture and observed under a stereomicroscope after harvest. Thirty and fifty plants were observed with mock and naringenin treatment, respectively. The ANOVA followed by Tukey's test showed no significant difference between mock treatment and naringenin addition in high moisture condition (p >0.05).

To further understand whether naringenin had effects on *Rl* Norway, we examined if naringenin impacted the *nodA* expression and root colonization of *Rl* Norway. The *nodA* gene encodes an N-acetyltransferase involved in Nod factor production and its expression can be induced by naringenin in *R. leguminosarum* bv. *viciae* (Tolin et al., 2013, Zaat et al., 1987). To investigate whether naringenin contributes to the symbiosis establishment between *L. japonicus* Gifu and *Rl* Norway via enhancing the transcription of *nod* genes, we inspected the activation of $nodA_{pro}$ by naringenin with an *Rl* Norway reporter strain (*Rl* Norway-$nodA_{pro}$:mCherry). This strain carried two plasmids of an mCherry under the control of $nodA_{pro}$ and a constitutively expressed cerulean fluorescent protein. The *Rl* Norway-$nodA_{pro}$:mCherry was cultured at 28˚C on plates for three days in combination with three concentrations of naringenin. In the presence of 0.1 µM naringenin, the cerulean fluorescence was observed while the mCherry signal was barely visible. The mCherry and cerulean fluorescence was observed in the presence of 1 µM and 10 µM naringenin (Figure 18). Naringenin activated $nodA_{pro}$ at 10 µM, followed by 1 µM. This result suggested a positive correlation between $nodA_{pro}$ and the concentration of naringenin. As naringenin had a growth inhibition effect on *Rl* Norway at 10 µM (Figure 17A), 1 µM naringenin was applied in the following experiment.

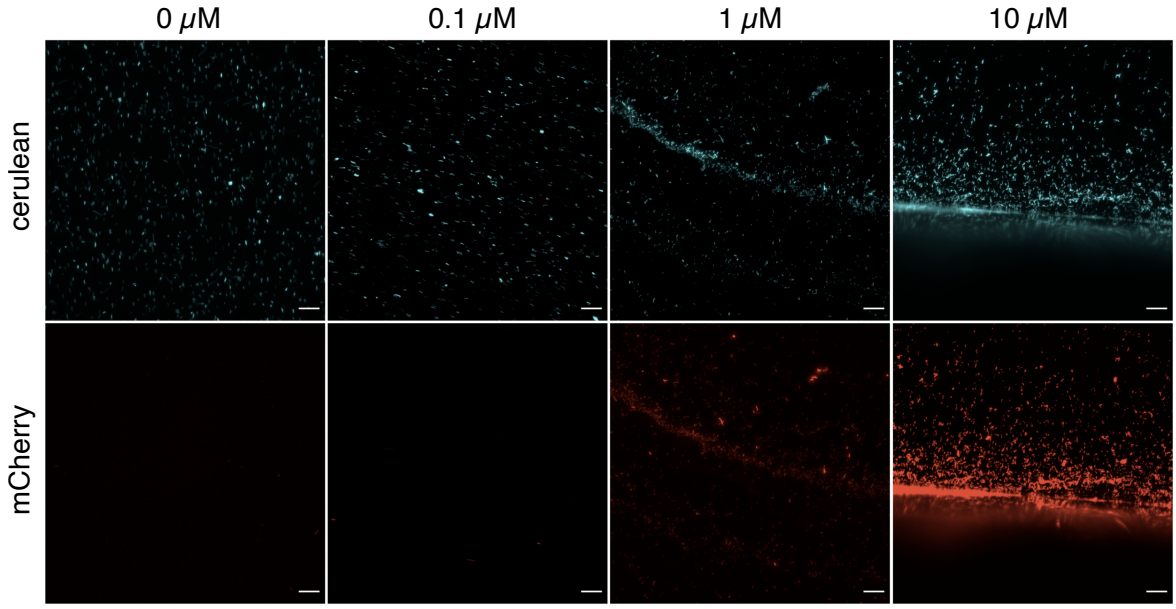


Figure 18. Induction *Rl* Norway *nodA* promoter by naringenin. A suspension of *Rl* Norway carrying the $nodA_{pro}$:mCherry and $SRKGm_{pro}$:cerulean plasmids was cultivated at 28˚C for 3 days on TY media plate supplemented with antibiotics and increasing concentrations of naringenin. The bacteria were examined under a fluorescence microscope. Scale bar = 25 µm.

To examine the response of *Rl* Norway-$nodA_{pro}$:mCherry *in planta* in the presence of 1 µM naringenin, *L. japonicus* Gifu plants grown in LM and HM were inoculated with *Rl* Norway-

nodA$_{pro}$:mCherry and the roots were observed at 2 wpi. In LM, *Rl* Norway barely colonized the root surface. The cerulean fluorescence indicated few rhizobia were dispersed at the root base and no mCherry signal was detected (Figure 19). On the contrary, *Rl* Norway distributed evenly from the root base to the root tip in HM. The mCherry signals were detected, but these signals did not co-localize with cerulean fluorescence, which indicated plasmid instability. Altogether, the results showed that naringenin induced *nodA* promoter *in planta*.

As the colonization of *Rl* Norway was strikingly high on *L. japonicus* Gifu grown in HM, we examined if naringenin treatment promoted rhizobia colonization, especially in LM. *L. japonicus* Gifu plants were inoculated with *Rl* Norway in the presence of 1 μM naringenin. The root colonization was quantified by qPCR at 2 wpi. For *L. japonicus* Gifu grown in LM, the colonization of *Rl* Norway was increased in naringenin-treated plants. A significant difference in the colonization between naringenin-treated and untreated groups was obtained (Figure 20). The *Rl* Norway colonization was nearly two times higher on roots grown in HM than in LM without naringenin treatment. The effect of naringenin on *Rl* Norway colonization was not significant in HM condition. This result suggested that naringenin can contribute to the *Rl* Norway colonization in LM.

In summary, genes categorized in the oxidation-reduction process were mainly influenced by substrate moisture after *Rl* Norway. Several genes involved in flavonoid biosynthesis were downregulated in HM, suggesting a role of flavonoids in high moisture-dependent nodulation. We examined whether naringenin contributes to the nodulation of *L. japonicus* Gifu with *Rl* Norway and found naringenin could induce the *nodA*$_{pro}$ of *Rl* Norway and increase *Rl* Norway colonization in LM.

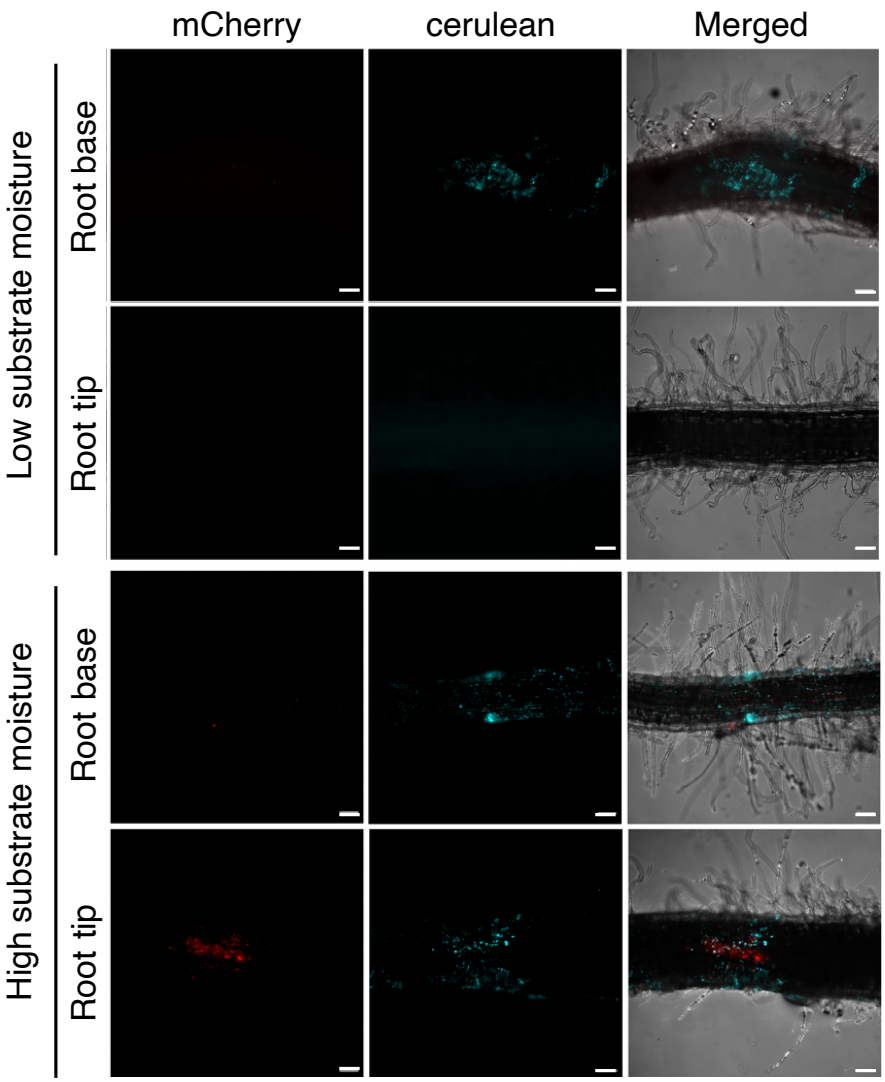


Figure 19. Induction of *Rl* Norway *nodA* promoter in high substrate moisture on *L. japonicus* Gifu root. Representative pictures of the *nodA* promoter activity at low and high substrate moisture conditions are shown. *Rl* Norway carrying the $nodA_{pro}$:mCherry and $SRKGm_{pro}$:cerulean plasmids was grown in TY medium supplemented with antibiotics and 1 µM naringenin for 2 days. Plants grown in low substrate moisture and high substrate moisture were inoculated with bacteria suspension (OD_{600} = 0.005) supplemented with 1 µM naringenin and harvested at 2 wpi. At least three roots from five different pots were examined under a fluorescence microscope. Scale bar = 100 µm.

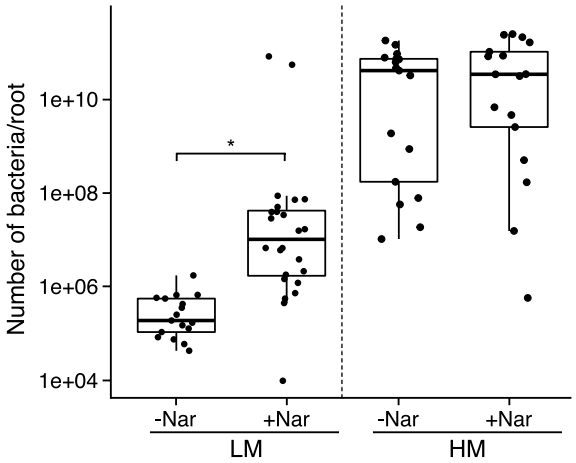


Figure 20. Quantification of root colonization by *Rl* Norway after naringenin treatment. *L. japonicus* Gifu were grown in low and high substrate moisture (LM and HM) and inoculated with *Rl* Norway (OD_{600} = 0.005) with or without 1 µM naringenin treatment (+Nar or -Nar). Root fragments (3 cm from the tip of the primary root) were harvested at 2 wpi from three different pots. Bacteria were detached from the root by incubating, sonicating, and vortexing in sterile water and quantified by qPCR. Each dot represents one individual plant. The asterisks indicate the significant difference determined by unpaired t-tests (p <0.01).

5 Investigation of candidate genes that contribute to the different symbiotic compatibility between *L. burttii* and *L. japonicus* Gifu

5.1 *L. burttii* and *L. japonicus* Gifu responded differently to *Rl* Norway

L. burttii nodulated with *Rl* Norway while *L. japonicus* Gifu did not in low substrate moisture condition. To investigate the response of *L. burttii* and *L. japonicus* Gifu to *Rl* Norway at the transcriptional level, transcriptomic profiles of *L. burttii* and *L. japonicus* Gifu grown in LM treated with *Rl* Norway and mock were investigated. In *L. burttii* and *L. japonicus* Gifu, DEGs were determined by comparing their transcriptome data of *Rl* Norway inoculated roots relative to their mock-treated roots using the DEseq2 package in R (Love et al., 2014). A total of 227 and 511 DEGs (p <0.01 and absolute value of $\log_2$ fold change >1) were determined in *L. burttii* and *L. japonicus* Gifu, respectively (Figure 21). In *L. burttii*, 227 genes were differentially regulated. Among them, 159 genes were upregulated and 68 genes were downregulated (Figure 21A). In *L. japonicus* Gifu, 10 of the DEGs were downregulated while 501 of them were upregulated. In both cases, the level of upregulation is higher than downregulation (Figure 21B).

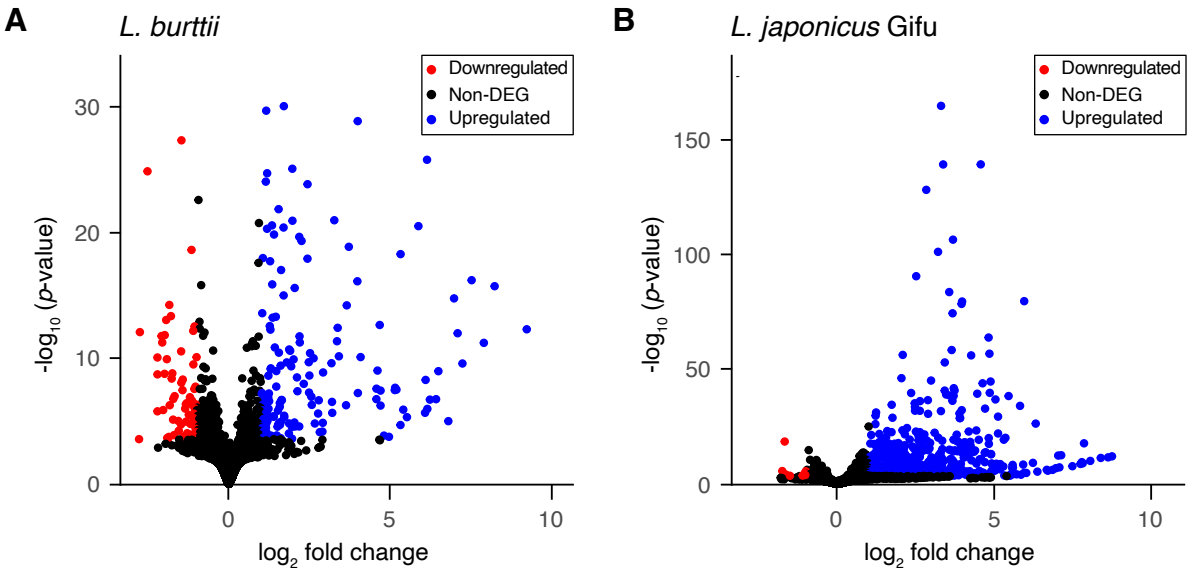


Figure 21. Genes differentially regulated in *L. burttii* and *L. japonicus* Gifu after *RI* Norway inoculation. Volcano plots of the relative expression level of genes regulated by *RI* Norway in two *Lotus* species. RNA-seq data were obtained from roots of *L. burttii* and *L. japonicus* Gifu grown in low substrate moisture condition for 14 days after *RI* Norway inoculation. Fold changes of the genes were relative to the mock condition. Genes with p <0.01 and absolute value of $\log_2$ fold change over 1 were considered as differentially regulated. The blue and red dots indicate upregulated and downregulated genes, respectively. Non-differentially expressed genes (Non-DEG) are depicted in black. The DEG analysis was conducted in R with the DESeq2 package (Love et al., 2014) and the volcano plot was produced with the ggplot2 package (Wickham, 2016).

Among the 227 DEG of *L. burttii*, 128 genes were functionally annotated. A total of 17 genes encoding enzymes with functions in cell wall organization or biogenesis, including xyloglucan endotransglucosylase/hydrolase, pectinesterase, xylosidase, hexosyltransferase, glycosyltransferase, and glucanase were identified (Le Gall et al., 2015). These enzymes are involved in polysaccharide metabolism, which are essential components for plant cell walls. For instance, five genes encoded xyloglucan endotransglucosylase/hydrolases that can digest and re-construct hemicellulose chains in the plant cell wall (Sharples et al., 2017). Enzymes processing pectin, a major component of the plant cell wall, function in cell adhesion (Daher & Braybrook, 2015). Three genes encoding expansin and expansin-like proteins were identified, which participate in cell expansion when cell-wall modification occurs (Sampedro & Cosgrove, 2005). In addition, three genes encoded *O*-methyltransferases and four genes encoded peroxidases were identified, which function in lignin biosynthesis and oxidation-reduction process correlated to cell-wall modification (Le Gall et al., 2015). Nodulin, nodulin-like, and early nodulin-like protein-encoding genes, which express during the establishment of nitrogen-fixing symbiosis (Bottomley & Myrold, 2007), were also identified. Other annotated genes were distributed in cellular and biological processes, including amino acid catabolic process,

cofactor catabolic process, and antibiotic catabolic process. In summary, the functional investigation into the DEG suggested *Rl* Norway inoculation activated the cell wall biosynthesis and metabolism in *L. burttii*.

In *L. japonicus* Gifu, nine senescence-associated proteins were highly expressed (log_2 fold change over 5) after *Rl* Norway inoculation, and a cytokinin oxidase/dehydrogenase-like protein was moderately upregulated (log_2 fold change around 2.8). The cytokinin oxidase/dehydrogenase functions in the maintenance of cytokinin homeostasis during root and nodule development (Reid et al., 2016), which may contribute to senescence by degrading cytokinin irreversibly (Galuszka et al., 2001; Kshishan and Rashotte, 2015). Genes encoding enzymes in cell wall modification (xyloglucan endotransglucosylase/hydrolases and subtilisin-like proteases), the TCP, Myb, basic HLH transcription factors, ribosomal proteins and, enzymes involved in the metabolic processes were differentially expressed mildly. In addition, genes encoding transporters and putative retrotransposon proteins were also differentially regulated. Conclusively, although *L. japonicus* Gifu did not nodulate with *Rl* Norway in LM, the DEG suggested the cellular metabolic processes were influenced.

To uncover whether the identified DEGs were correlated with symbiosis response, we investigated the expression pattern of the DEG in a broad range of *Lotus* transcriptomic profiles. Available transcriptome datasets from plant organs (seed, pod, mature flower, immature flower, leaf, shoot, and root), root hairs after inoculation conditions, and either pathogenic bacteria (*R. solanacearum* JS763, *P. syringae* pv. tomato DC3000) or symbiotic bacteria (*Ef* HH103, *B. elkanii* USDA61, and *M. loti* R7A) inoculated roots were acquired from Lotus Base (https://lotus.au.dk/). The top 25 DEGs of *L. burttii* and *L. japonicus* Gifu were used as queries to obtain their expression profiles, respectively. The expression of the top 25 DEGs in *L. burttii* showed a positive correlation with *M. loti* R7A inoculation, which is the compatible symbiont of *Lotus*. Those DEGs possessed a higher expression in the transcriptome of root hairs, roots, nodule primordia, and nodules after *M. loti* R7A inoculation (Figure 22). Most of the DEGs were barely expressed in seed, pod, mature flower, immature flower, leaf, and shoot as well as in the roots inoculated with pathogenic bacteria and non-compatible symbiotic rhizobia. Two exceptions were genes encoding a peroxidase and a short-chain dehydrogenase/reductase. The former was highly expressed in pods and the latter was highly expressed with either pathogenic or symbiotic bacteria inoculation. In contrast, the expression pattern of DEGs in *L. japonicus* Gifu was more dispersed. Generally, the DEG exhibited a relatively higher expression either in root and root hairs regardless of the type of inoculum or in plant organs (Figure 23). However, two senescence-associated protein-encoding genes, a glycosyltransferase, a cytochrome P450 family protein and a BURP protein encoding gene

were lowly expressed in most conditions. Overall, the positive correlation between the expression pattern of the DEGs and compatible symbiont inoculation supported that a symbiotic response was induced in *L. burttii*. The dispersed expression pattern of DEG in *L. japonicus* Gifu suggested that *Rl* Norway either did not trigger the symbiosis-specific response or the triggered response was undetectable by RNA-seq.

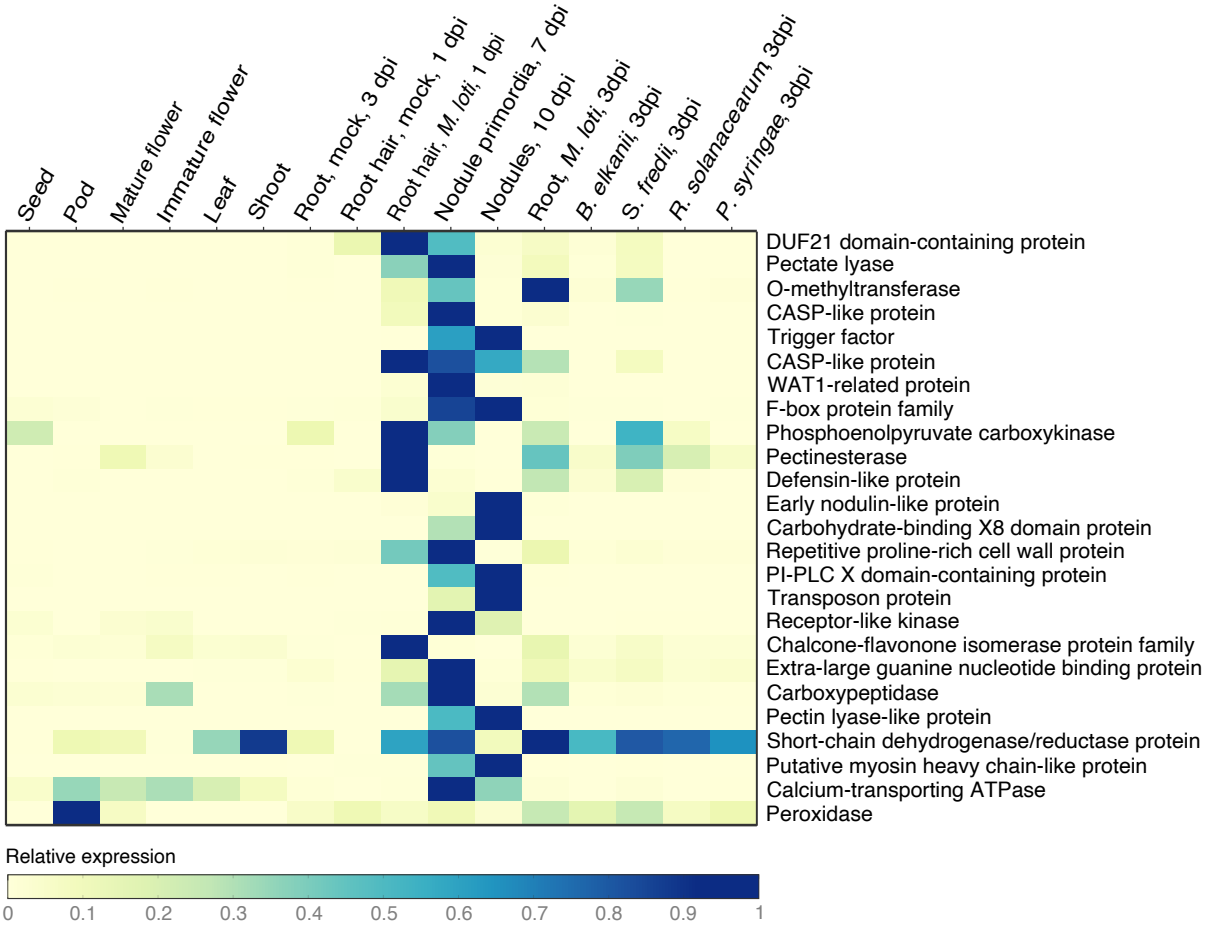


Figure 22. Expression pattern of the top 25 differentially expressed genes in *L. burttii* under symbiotic and non-symbiotic conditions. A total of 16 transcriptome datasets were obtained from Lotus Base (https://lotus.au.dk/) to examine the expression pattern of the top 25 differentially expressed genes (DEGs) in *L. burttii*. The top 25 DEGs of *L. burttii* were used as queries to acquire the expression pattern. The expression level of the DEGs was normalized across conditions.

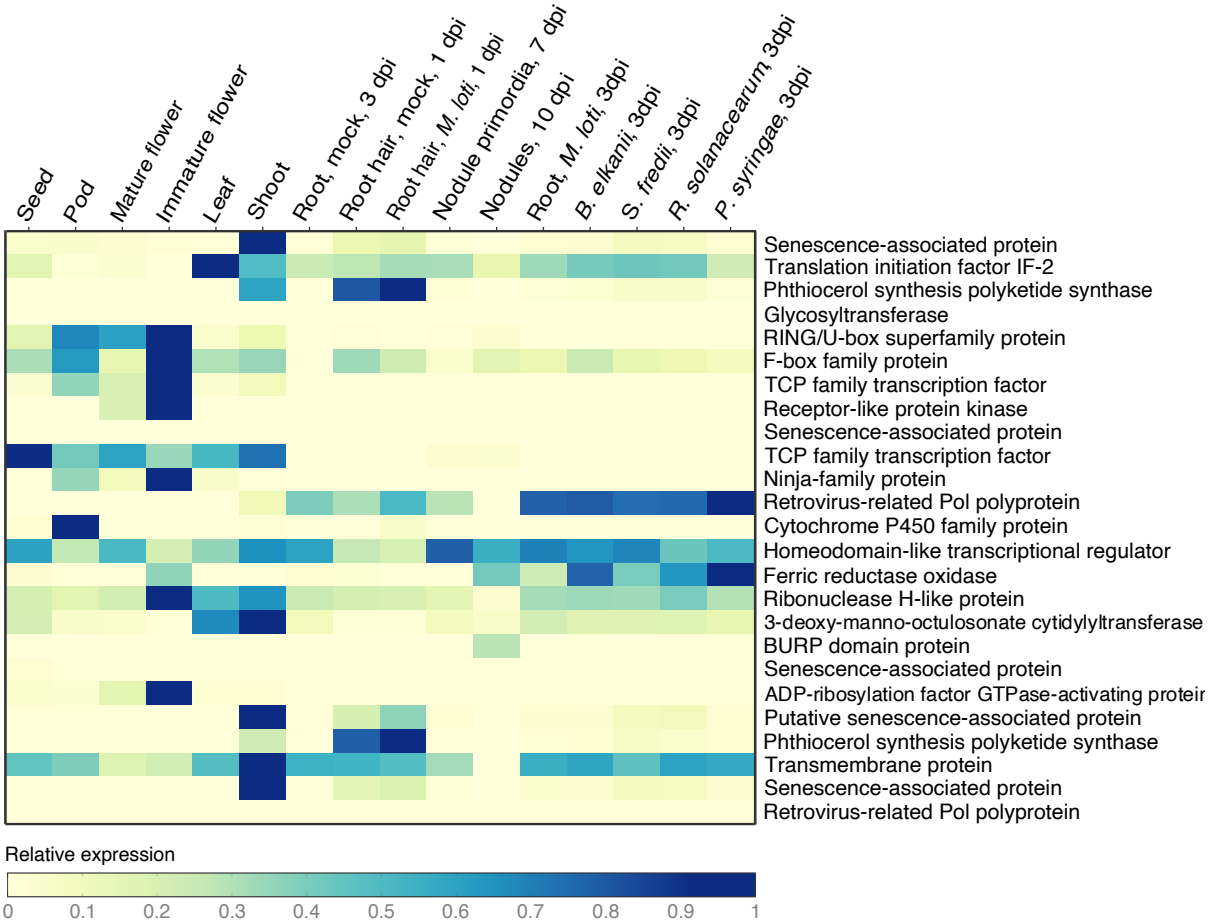


Figure 23. Expression pattern of the top 25 differentially expressed genes in *L. japonicus* Gifu under symbiotic and non-symbiotic conditions. A total of 16 transcriptome datasets were obtained from Lotus Base (https://lotus.au.dk/) to examine the expression pattern of the top 22 differentially expressed genes (DEGs) in *L. japonicus* Gifu. The top 25 DEGs of *L. japonicus* Gifu were used as queries to acquire the expression pattern. The expression level of the DEGs was normalized across conditions.

5.2 Symbiotic response was activated in *L. burttii* but not in *L. japonicus* Gifu in response to *RI* Norway

As *L. burttii* nodulated with *RI* Norway and DEG suggested that *L. burttii* possessed a symbiosis-specific response but *L. japonicus* Gifu did not, we inspected the expression of the symbiosis marker genes via a targeted approach. At 2 wpi, *Nodule inception* (*NIN*) and *exopolysaccharide receptor 3* (*EPR3*) were upregulated over 4-fold in *L. burttii* (Figure 24). *Nodulation pectate lyase* (*NPL*), which is induced by Nod factors via activation of the nodulation signaling pathway and the NIN transcription factor, was significantly upregulated especially in *L. burttii*. In addition, *SymRK-interacting E3 ubiquitin ligase* (*SIE3*) and *NENA* were slightly upregulated in *L. burttii* but downregulated in *L. japonicus* Gifu. Moreover, the *ethylene response factor required for nodulation 1* (*ERN1*) was upregulated in *L. burttii*, which relates

to rhizobia infection. This indicated that early symbiotic genes were induced in *L. burttii* but not in *L. japonicus* Gifu, suggesting that *L. burttii* activated a symbiotic response to *Rl* Norway.

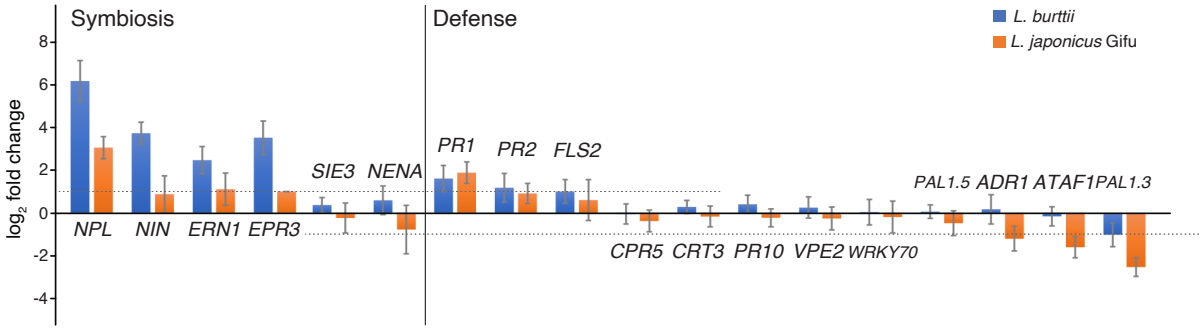


Figure 24. The expression level of symbiosis and defense marker genes in *L. burttii* and *L. japonicus* Gifu. The bar plot shows the fold-change of genes after *Rl* Norway inoculation relative to the mock condition. These genes are categorized by function according to previous studies (Bhardwaj et al., 2011; Hara-Nishimura et al.,2005; Roy et al.,2020; Wang et al. 2016). *NIN, nodule Inception*; *SIE3, SymRK-interacting E3 ubiquitin ligase*; *EPR3*, *exopolysaccharide receptor 3*; *NPL, Nodulation pectate lyase*; *ERN1, ethylene response factor required for nodulation 1*. Together with *NENA*, these genes are involved in symbiosis response and establishment. *Pathogenesis-related gene 1* and *2* (*PR1* and *PR2*) and *constitutive expression of PR genes 5* (*CPR5*) are involved in systemic acquired resistance. *Flagellin-sensing 2* (*FLS2*) and *Calreticulin 3* (*CRT3*) are bacteria-responding genes. *Class 10 pathogenesis-related gene* (*PR10*), *WRKY transcription factor 70* (*WRKY70*), *Phenylalanine ammonia-lyase* (*PAL*), and *Activated disease resistance 1* (*ADR1*) are involved in salicylic acid-dependent defense. *Vacuole processing enzyme* 2 (*VPE2*) and *Arabidopsis transcription activation factor 1-like protein* (*ATAF1*) are involved in hypersensitive response and abscisic acid response, respectively. The dash lines indicate the 2 and 0.5 fold change in expression (log_2 = 1 and -1).

Different from *L. burttii*, the transcriptomic data suggested that the symbiotic response was not induced in *L. japonicus* Gifu. To understand whether *Rl* Norway triggered a defense response instead in *L. japonicus* Gifu, the expression of defense marker genes and genes reacting to bacterial contact were examined via a targeted approach. In general, no significant difference was found in the expression level of molecular markers of systemic acquired resistance (*Pathogenesis-related gene 1* and *2*), salicylic acid-dependent defense genes (*Class 10 pathogenesis-related gene*, *WRKY transcription factor 70*, and *phenylalanine ammonia-lyase 1.5*), and a bacterial flagellin-responding gene (*Flagellin-sensing 2*) between the two *Lotus* species (Figure 24). None of the inspected genes was strongly upregulated. The change in expression of several genes was lower than 2 fold or barely detectable. Only three of the inspected genes were downregulated over 2 fold in *L. japonicus* Gifu. In summary, the expression level of most inspected defense-related genes was similar in *L. burttii* and *L. japonicus*. These data suggested that the level of defense response was not significantly different between *L. burttii* and *L. japonicus* Gifu at given condition, indicating the defense

response may not be a major contributor to the different symbiotic compatibility between the two *Lotus* species.

5.3 Using genetic approaches to identify genes contributing to the difference in symbiotic compatibility between *L. burttii* and *L. japonicus* Gifu

5.3.1 Twelve candidate genes were identified in a QTL that co-segregated with the nodulation phenotype of *L. burttii*

To determine if the nodulation phenotype in *L. burttii* is a dominant trait, 26 F1 plants of a cross between *L. burttii* and *L. japonicus* Gifu were inoculated with *Rl* Norway and phenotyped. The F1 progeny comprised plants generated either using *L. burttii* as maternal plant and *L. japonicus* Gifu as paternal plant or vice versa. Plants were genotyped by PCR using *Lotus* power marker TM1203 (Figure 25A). *L. burttii* and the F1 plants, regardless of their parental combination, nodulated with *Rl* Norway at 6 wpi, although the number of nodules and nodule primordia varied. *L. japonicus* Gifu did not nodulate at all (Figure 25B). These results indicated that the nodulation phenotype of *L. burttii* was dominant.

To find what genes control the symbiotic compatibility, a QTL mapping on 146 recombinant inbred lines (RILs) of *L. burttii* and *L. japonicus* Gifu was conducted. A 0.3 Mb region co-segregated with the nodulation phenotype (M. Parniske and S. U. Andersen, personal communication) and was located on the upper arm of chromosome 1 in *L. japonicus* Gifu. To investigate if this QTL is conserved in legumes, a synteny analysis was performed. The QTL of *L. japonicus* Gifu was used as a query to search the syntenic region in *L. japonicus* MG-20, *Medicago truncatula* A17, and *Phaseolus vulgaris* BAT93. The synteny analysis showed that this QTL was conserved among the examined species (Figure 26), suggesting a conserved role of this region in legumes.

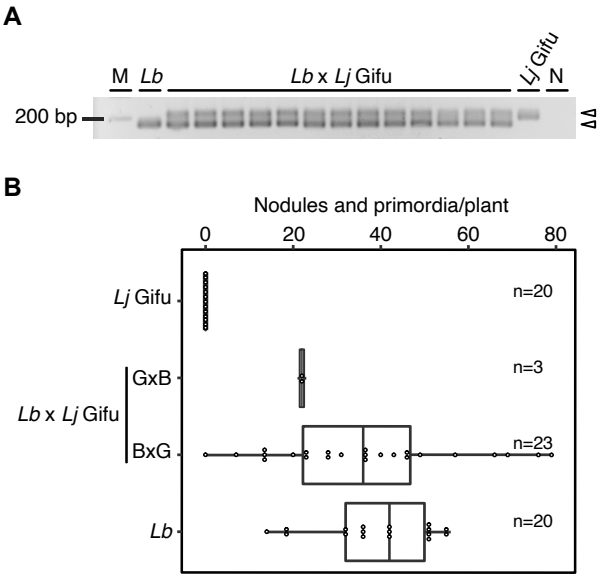


Figure 25. The nodulation phenotype of *L. burttii* is dominant. Twenty-six F1 plants of a cross between *L. burttii* and *L. japonicus* Gifu were inoculated with *Rl* Norway (OD_{600} = 0.005) and examined at 6 wpi. (A) Genotyping of the F1 progeny. Genomic DNA was extracted from one leaf of each plant, genotyped by PCR with *Lotus* power marker TM1203, and analyzed by acrylamide gel electrophoresis. Genotyping results of 13 randomly selected plants are presented. M, DNA ladder; N, negative control. The upper and lower arrowheads indicate the amplification of markers from *L. japonicus* Gifu and *L. burttii*, respectively. (B) Nodulation phenotype of the F1 progeny. Progeny with *L. japonicus* Gifu as the mother plant and *L. burttii* as pollen donor is indicated as GxB. Progeny with *L. burttii* as the mother plant and *L. japonicus* Gifu as pollen donor is indicated as BxG. *Lb*, *L. burttii*; *Lj* Gifu, *L. japonicus* Gifu.

To identify candidate genes that are located in the QTL and contribute to the difference in nodulation compatibility, the genomes of *L. japonicus* Gifu and *L. burttii* (shared by S.U. Andersen) were compared. Due to error-prone automatic genome annotation, we manually re-annotated the genome of *L. burttii* and *L. japonicus* Gifu. Around 40 genes were predicted. Single nucleotide polymorphisms (SNPs) between the two *Lotus* accessions were identified using CLC Main Workbench 7 (Qiagen). Roughly 2,700 SNPs were distributed in the QTL. The QTL was comprised of hypervariable regions where many SNPs were found in close proximity and hypovariable regions where no or few SNPs were detected. Most of the SNPs were located in intergenic regions, while around 520 SNPs were located in introns and 430 were found in exons. 97% of the SNPs in coding sequences (CDSs) encoded silent or missense mutations, while 3% of the SNPs led to frameshifts or nonsense mutations. Twelve genes showed polymorphisms that led to early stop codons in CDS between the two *Lotus* accessions (Figure 27). The 12 candidates were numbered according to their relative position on the QTL (Figure 27A). The SNPs in the 12 genes caused mostly missense mutations, while few of them encoded nonsense or frameshift mutations. Six of the genes had premature stop codons in

L. burttii, while the other six genes had premature stop codons in *L. japonicus* Gifu (Figure 27B). The 12 candidates were predicted to encode four types of proteins: proteins with unknown function, proteins associated with transposable elements (TE), leucine-rich repeat (LRR) containing resistance proteins, and LRR containing receptor-like proteins (RLPs).

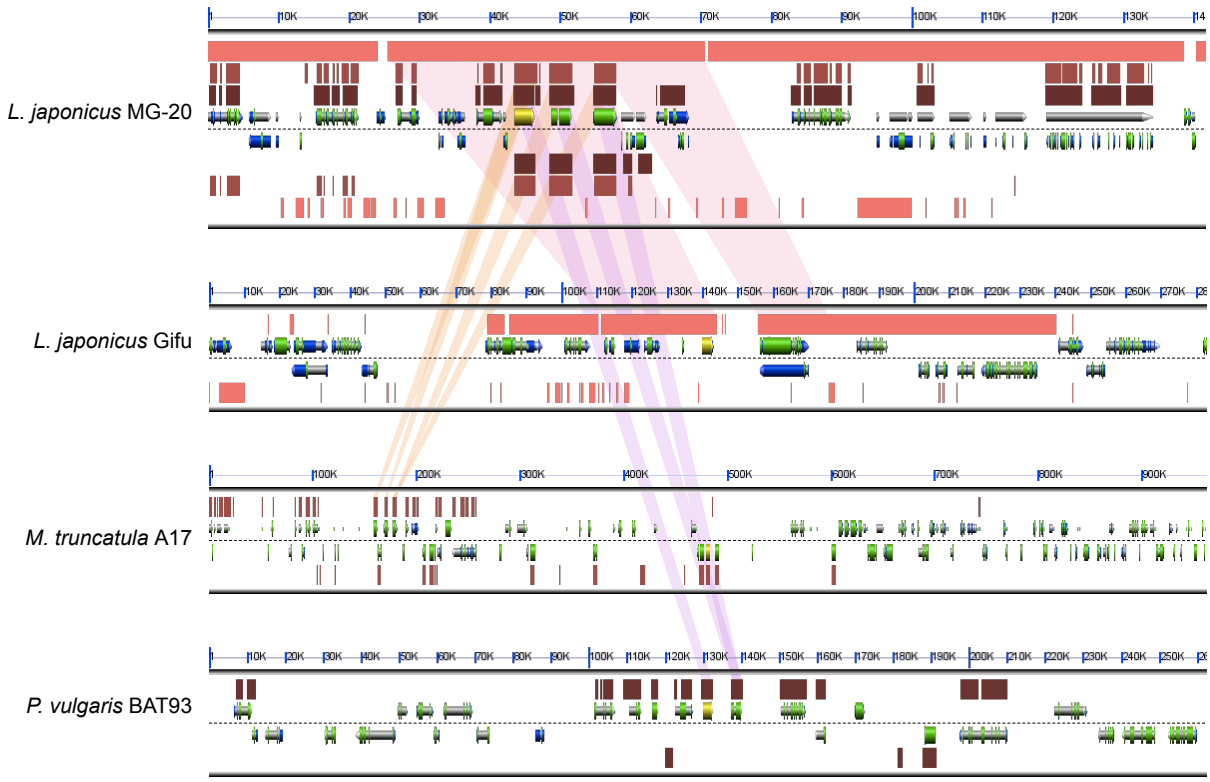
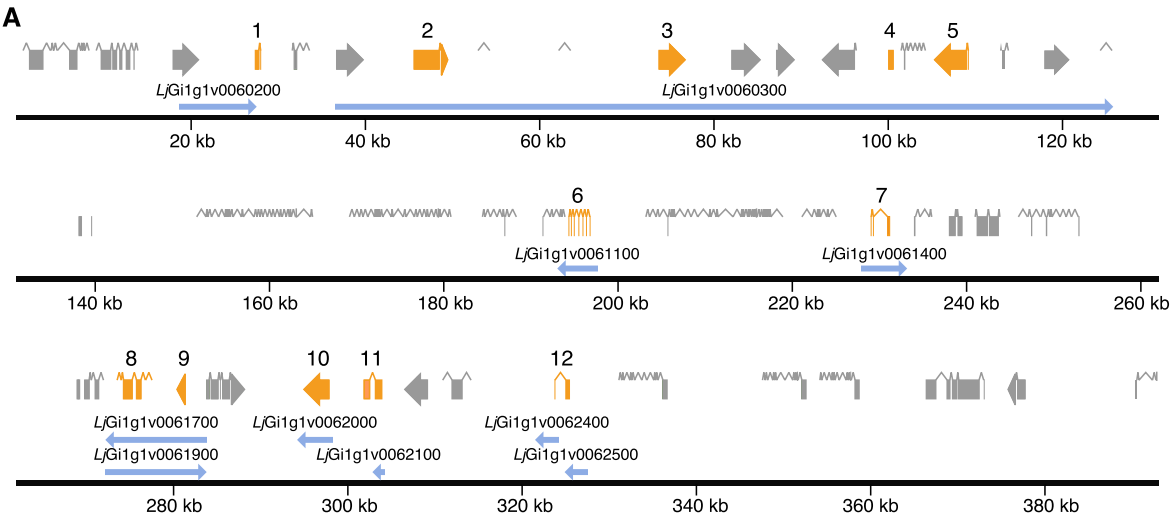


Figure 26. Synteny analysis of the *RLP* cluster in *L. japonicus*, *M. truncatula*, and *P. vulgaris*. The synteny analysis was conducted on CoGe: SynFind platform (Lyons and Freeling, 2008) with the comparison algorithm Last and specified feature Lj2g3v2904830.1 (*RLP1*, depicted in yellow). The gene models annotated in the database are presented in the middle of each panel, in combinations of the colors grey, green and blue. The color blocks above and underneath the gene models are homologous genes located on the positive strand and the negative strand, respectively. The color wedges indicate the homologous genes between *L. japonicus* MG-20 and *L. japonicus* Gifu, *M. truncatula* A17, or *P. vulgaris* BAT93.

Candidates 6, 7, and 12 encoded proteins with unknown functions. The proteins encoded by candidates 6 and 12 were annotated as CM0545.280.nc and At3g47200-like protein, respectively. The former was truncated in *L. burttii* and the latter was *L. japonicus* Gifu. In both cases, the truncated proteins were more than 200 amino acids shorter compared to the predicted full-length ones. In contrast, the putative gene product of *LOC100797891* (candidate 7) was only 8 amino acids shorter in *L. japonicus* Gifu than in *L. burttii*.

A



B

Annotated function of candidate genes	Location of SNPs		Mutations			
	Intron	Exon	Silent	Missense	Nonsense	Frameshift
1. Ribonuclease H[1]	1	4	1	2	0	1
2. Putative disease resistance protein RGA4[2]	9	46	8	37	1	0
3. Retrotransposon protein[2]	0	9	1	7	1	0
4. NBS-LRR resistance protein[1]	0	5	2	2	0	1
5. Ribonuclease H[2]	1	18	2	14	1	1
6. CM0545.280.nc[2]	2	5	1	4	0	1
7. Uncharacterized protein LOC100797891[1]	1	1	0	0	1	0
8. Hypothetical protein VITISV_01845[2]	2	17	3	12	0	1
9. RLP4[1]	0	5	0	2	0	1
10. RLP3[2]	0	2	0	1	1	0
11. RLP2[1]	6	3	2	1	0	1
12. UPF0481 protein At3g47200-like/ DUF247 domain containing protein[1]	0	8	4	3	1	0

[1]Mutations leading to early stop codon in *Lj* Gifu
[2]Mutations leading to early stop codon in *L. burttii*

Figure 27. Twelve candidate genes were identified in the QTL region that co-segregated with the nodulation phenotype. (A) Gene structures and relative positions in the *NFR1*-linked region. The CDS of the twelve candidate genes is numbered and depicted in orange. The CDS of non-candidate genes is depicted in grey. All CDSs are depicted in reference to *L. japonicus* Gifu. The gene loci are depicted in blue with their gene loci ID above the blue arrows. The direction of the arrows indicates the direction of the gene loci. The relative position of the genes is indicated below in kilobases (kb). *Lj*Gi: LotjaGi, the gene loci ID in *L. japonicus* Gifu genome v1.2 (Kamal et al., 2020). (B) The annotated function of the candidate genes. The table presents the predicted function, the number of SNPs, and the number of amino acid changes of each candidate gene. The CDSs were predicted by FGENESH (Solovyev et al., 2006) and annotated manually by NCBI blastx (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The candidate genes were chosen based on structural differences and the existence of premature stop codon in the encoding protein between *L. burttii* and *L. japonicus* Gifu predicted by InterPro webtool (Hunter et al., 2009).

A retrotransposon (candidate 3) and a hypothetical protein VITISV_01845 (candidate 8) encoding genes were identified to be truncated in *L. burttii*. In the hypothetical protein VITISV_01845, an integrase domain, a ribonuclease domain, a gag-pol domain, and a copia-like domain were predicted. Those domains were typically present in transposable elements (Sabot and Schulman 2006). Candidate 1 and 5 encoded proteins containing a ribonuclease H domain, which is often part of retrotransposons and mediates the cleavage of the RNA strand of an RNA–DNA duplex (Ustyantsev et al, 2015). While candidate 1 contained an early stop codon in *L. japonicus* Gifu, candidate 5 contained an early stop codon in *L. burttii*. Putative proteins encoded by candidate genes in this group displayed size differences of more than 400 amino acids compared to the respective full-length protein. Moreover, the gene loci where candidate 1 to 5 were located contained several repetitive sequence pieces.

Two proteins belonging to the nuclear-binding site-leucine-rich repeat (NBS-LRR) family were encoded by candidate 2 and candidate 4 and were annotated as 'resistance gene analogue 4' (RGA4) and 'NBS-LRR', respectively. The two NBS-LRR containing proteins had striking differences in length. The former was predicted as a full-length protein containing 1,305 amino acids in *L. japonicus* Gifu and was 105 amino acids shorter in *L. burttii.* The latter was 219 and 233 amino acids in *L. japonicus* Gifu and *L. burttii*, respectively. The candidate 4 encoding NBS-LRR resistance protein was predicted to be truncated in both *Lotus* species.

Four genes encoding LRR containing receptor-like proteins (RLP) were found in the QTL. They were located proximally and named *RLP1*, *RLP2* (candidate 11), *RLP3* (candidate 10), and *RLP4* (candidate 9). The protein encoded by *RLP1* is highly conserved between *L. burttii* and *L. japonicus* Gifu with 99.8% identity, whereas the proteins encoded by RLP2, RLP3, and RLP4 comprised SNPs that lead to early stop codons. The truncated RLPs were 40% to 80% shorter than the respective full-length form. RLP2 and RLP4 were truncated in *L. japonicus* Gifu, whereas RLP3 was truncated in both accessions. The RLP2 and RLP4 consisted of 558 and 368 residues in *L. japonicus* Gifu (*Lj*RLP2 and *Lj*RLP4), while that in *L. burttii* were 924 and 983 amino acids in length (*Lb*RLP2 and *Lb*RLP4). RLP3 in *L. burttii* comprised only 184 amino acids (*Lb*RLP3), while that in *L. japonicus* Gifu contained 1,011 amino acids (*Lj*RLP3). Structure-wise, the RLPs were predicted to have the transmembrane domain and cytoplasmic tail in the full-length protein. RLP1, RLP3, and RLP4 were predicted to contain a canonical signal peptide at N-terminus, which was absent in RLP2. RLP3 in both *Lotus* species missed the transmembrane domain and the cytoplasmic tail.

5.3.2 Investigating the expression pattern of the twelve candidate genes

To see if the candidate genes were expressed in *L. burttii* and *L. japonicus* Gifu, we investigated their expression level in our transcriptomic data of *L. burttii* and *L. japonicus* Gifu under mock and *RI* Norway inoculation conditions. The ID of each gene locus, where the candidates were located (Figure 27A), was used to extract the number of reads mapped against each gene locus. Most of the candidate genes possessed read counts less than ten and were not significantly different between conditions. One exception was LotjaGi1g1v0060300, where candidates 2, 3, 4, and 5 were located (Figure 28). The read counts of LotjaGi1g1v0060300 gene were significantly greater than the read counts of all other gene loci regardless of the plant genotype (*L. japonicus* Gifu or *L. burttii*) and the experimental condition (mock treatment or *RI* Norway inoculation). Read counts of LotjaGi1g1v0060300 were higher in mock conditions than in *RI* Norway-inoculated conditions, especially in *L. japonicus* Gifu. However, the significant difference was found only between *L. burttii* and *L. japonicus* Gifu under mock condition ($p = 0.0095$). These results indicated that the genes may be either not expressed or lowly expressed and undetectable by RNAseq at the examined time point, suggesting that additional criterium should be incorporated to narrow down the candidate genes for further investigation.

As the nodulation phenotype of *L. burttii* was dominant and defense genes were not induced in *L. japonicus* Gifu upon *RI* Norway inoculation, we hypothesized that the genes controlling nodulation compatibility would be dominant rather than dominant negative. Therefore, we further inspected genes that were full-length in *L. burttii* and truncated in *L. japonicus* Gifu. Five candidate genes matching this criterium were candidates 1, 7, 9, 11, and 12, encoding a ribonuclease H, an uncharacterized protein, RLP4, RLP2, and a protein with unknown function. Since candidate 1 was annotated in the intergenic region without gene loci annotated in the *L. japonicus* Gifu reference genome and located in the highly repetitive region together with candidate genes encoding a retrotransposon protein and another ribonuclease H, it was not considered as a primary candidate to be investigated. As the genome of *L. burttii* was a draft genome based on Illumina sequencing data and some of the candidates were not well covered by the reads, we sequenced the region where candidates 7, 9, 11, and 12 located, to confirm the existence of the identified SNPs. The sequencing result showed that the intergenic region flanking candidate 12 was about 10 kb shorter in *L. burttii* (4,619 bp) than in *L. japonicus* Gifu (15,024 bp). Altogether, the RLP2 and RLP4 encoding genes were prioritized for advanced investigation, followed by the RLP3 encoding gene as RLPs have been reported to function in plant growth and development as well as symbiosis.

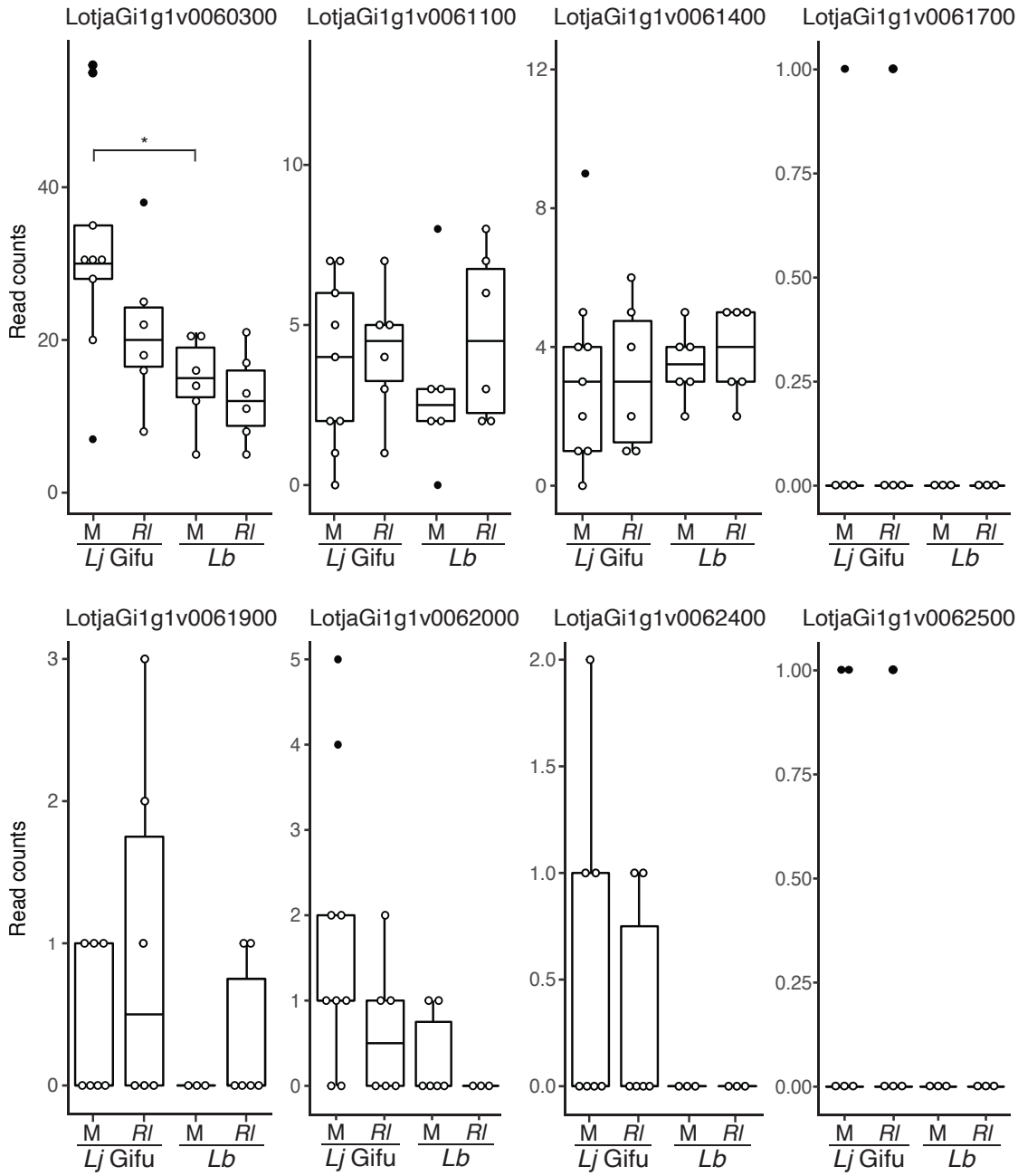


Figure 28. The expression of the candidate genes in the transcriptomes of *L. burttii* and *L. japonicus* Gifu. Transcriptomes of roots of *L. japonicus* Gifu (*Lj* Gifu) and *L. burttii* (*Lb*) at 2 wpi with *RI* Norway (*RI*) or mock treatments (M) grown in sand-vermiculite supplemented with 50 ml FAB medium were generated. The read count assigned to each locus is presented. For *L. burttii*, six root transcriptomes of each condition were analyzed. For *L. japonicus* Gifu, nine and six root transcriptomes with mock treatment and *RI* Norway inoculation were analyzed, respectively. The black dots indicate the outliers. The asterisk indicates the significant difference determined by a two-tailed Welch t-test (p <0.05).

As not all *RLP*s were found in our transcriptomic data, primers specific to each RLP were designed and a qRT-PCR was performed. The expression of *RLP*s was examined at 2 wpi in *L. burttii* and *L. japonicus* Gifu under mock and *RI* Norway-inoculated conditions. At 2 wpi,

RLP1 was barely expressed in *L.* burttii and *L. japonicus* under both conditions. Variations were found in the expression of *RLP2* and *RLP3*, but no significance was identified. Although the expression of *RLP4* was low, the expression of *LbRLP4* and *LjRLP4* RLP4 were higher in *Rl* Norway-inoculated condition than in mock condition (Figure 29).

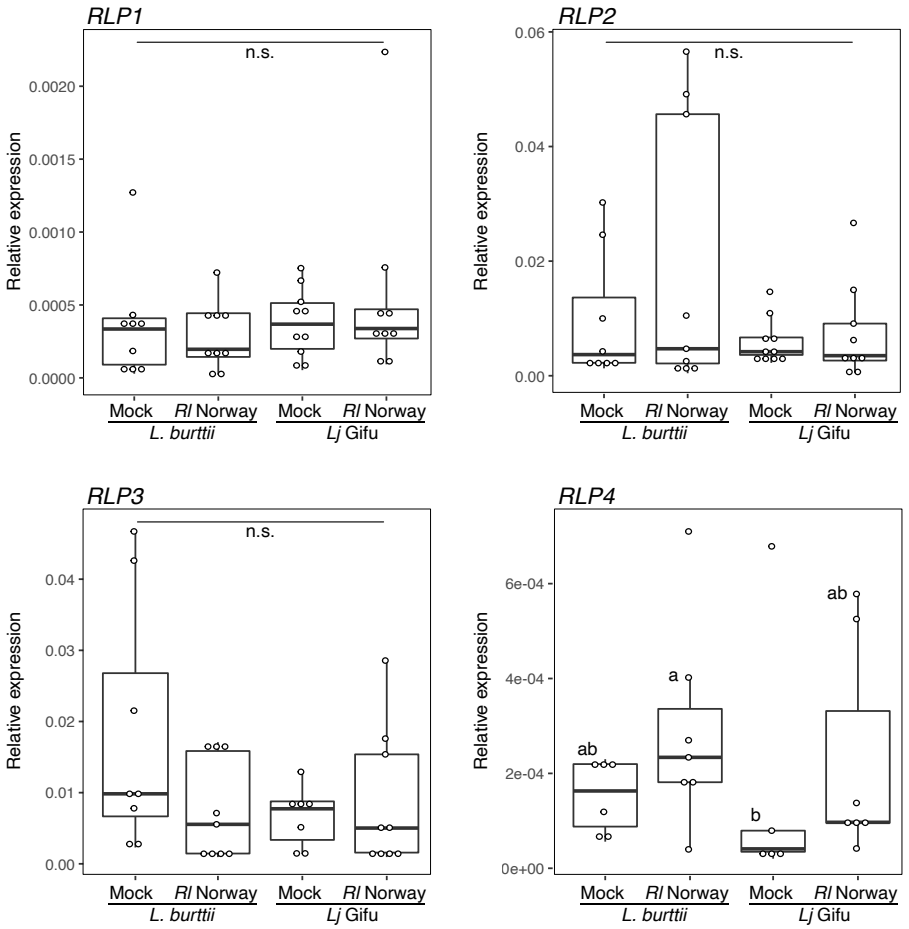


Figure 29. Expression analysis of *RLP*s in *L. burttii* and *L. japonicus* Gifu (*Lj* Gifu). The expression of *RLP1*, *RLP2*, *RLP3*, and *RLP4* was determined by qRT-PCR at 2 wpi. Ten biological replicates were used to perform the qRT-PCR. The relative expression of *RLP*s was normalized to *ATPsynthase*. Different lowercase letters indicate significant differences determined by ANOVA followed by Tukey's test (p <0.05).

5.3.3 Investigating the role of *RLP*s in symbiotic compatibility via *rlp*-mutant lines

To examine if the *RLP*s have a role in nodulation compatibility, the nodulation of *rlp*-mutant lines was inspected. As previous evidence indicated that *L. japonicus* Gifu could nodulate with *Rl* Norway in HM but not LM, we aimed to know whether the plants can nodulate with *Rl* Norway in LM or lose nodulation in HM when each *RLP* was mutated. The *rlp*-mutant lines in *L. japonicus* Gifu were insertion mutants generated by an endogenous long terminal repeat (LTR) retrotransposon *Lotus* Retrotransposon 1 (LORE1). At the moment when experiments were

designed, the *rlp4*-mutant lines were not available. Therefore, four *rlp1*-, six *rlp2*- and six *rlp3*-mutant lines were ordered from Lotus Base (https://lotus.au.dk/lore1/order). Homozygous mutants in *RLP*s and wild-type in *RLP*s with potential mutations in the background were screened and propagated for seed production. The mutant lines were phenotyped with *Rl* Norway and the phenotyping results were presented (Figure 30). At 6 wpi, no nodulation can be observed on plants grown in LM (data not shown), whereas plants consistently nodulated with *Rl* Norway in HM regardless of their genotype (Figure 30B to G). The nodulation ability is indistinguishable between wild type and *rlp1* plants (Figure 30B). Although significant differences in the number of nodules and nodule primordia were found between wild type and *rlp2*- or *rlp3*-mutants (Figure 30C, E to F), the results between different mutant lines within the same *RLP* was not consistent with each other. This may result from the different background mutations in the LORE1 lines.

5.3.4 *L. japonicus* Gifu plants trans-complemented with *LbRLP2* and *LbRLP4* partially nodulate with *Rl* Norway

*Lb*RLP2 and *Lb*RLP4 were predicted to be full-length in *L. burttii*, whereas *Lb*RLP3 was truncated in both *L. burttii* and *L. japonicus* Gifu (Figure 31A). To investigate if the *RLP* genes were responsible for the difference in nodulation phenotype between *L. burttii* and *L. japonicus* Gifu, we modified the *RLP* genotypes in *L. japonicus* Gifu to mimic those in *L. burttii* by trans-complementation assays. To this end, *LbRLP2* and/or *LbRLP4* under the control of an over-expression promoter (*oxLbRLP2* and *oxLbRLP4*) or their respective native promoters (*npLbRLP2* and *npLbRLP4*) were trans-complemented into *L. japonicus* Gifu and *rlp3-2* via *Agrobacteria*-mediated hairy root transformation. Transformed plants were further inoculated with *Rl* Norway. To clarify if *LbRLP2* and/or *LbRLP4* can confer the nodulation to *L. japonicus* Gifu, transformed plants were grown in LM after *Rl* Norway inoculation. On average, 20% of the *L. japonicus* Gifu and *rlp3-2* trans-complemented with *oxLbRLP2* nodulated at 8 wpi (Figure 31B). About 50% of the *L. japonicus* Gifu trans-complemented with ox*LbRLP4* and *oxLbRLP2+4* nodulated. About 33% of the *L. japonicus* Gifu and *rlp3-2* trans-complemented with np*LbRLP2+4* nodulated, while the other combinations showed a nodulation rate around 20%. Additionally, the nodulation events in *rlp3*-2 were lower than that in *L. japonicus* Gifu, although the difference was not significant. In general, plants trans-complemented with *LbRLP4* and *LbRLP2+4* tended to generate more nodules and nodule primordia regardless of the controlling promoter (Figure 31B).

As the nodulation phenotype did not meet normality, the Kruskal-Wallis test followed by Dunn's post-hoc-test was used to analyze the data. *L. japonicus* Gifu with *oxLbRLP2*, *oxLbRLP2+4*, and *npLbRLP2* showed significant differences compared with the control group (TM). In *rlp3*-

2, only plants with *npLbRLP2+4* were significantly different from the control group (Figure 31C). Nevertheless, an increasing tendency was observed. The nodulation increased progressively when trans-complemented *LbRLP2*, *LbRLP4*, and *LbRLP2+4* independent of the controlling promoter and plant genotype. Moreover, the statistical analysis also indicated a significant difference between *LbRLP2* and *LbRLP4* as well as between *LbRLP2* and *LbRLP2+4* trans-complementation. To conclude, *LbRLP2* and *LbRLP4* may have an additive effect on nodulation.

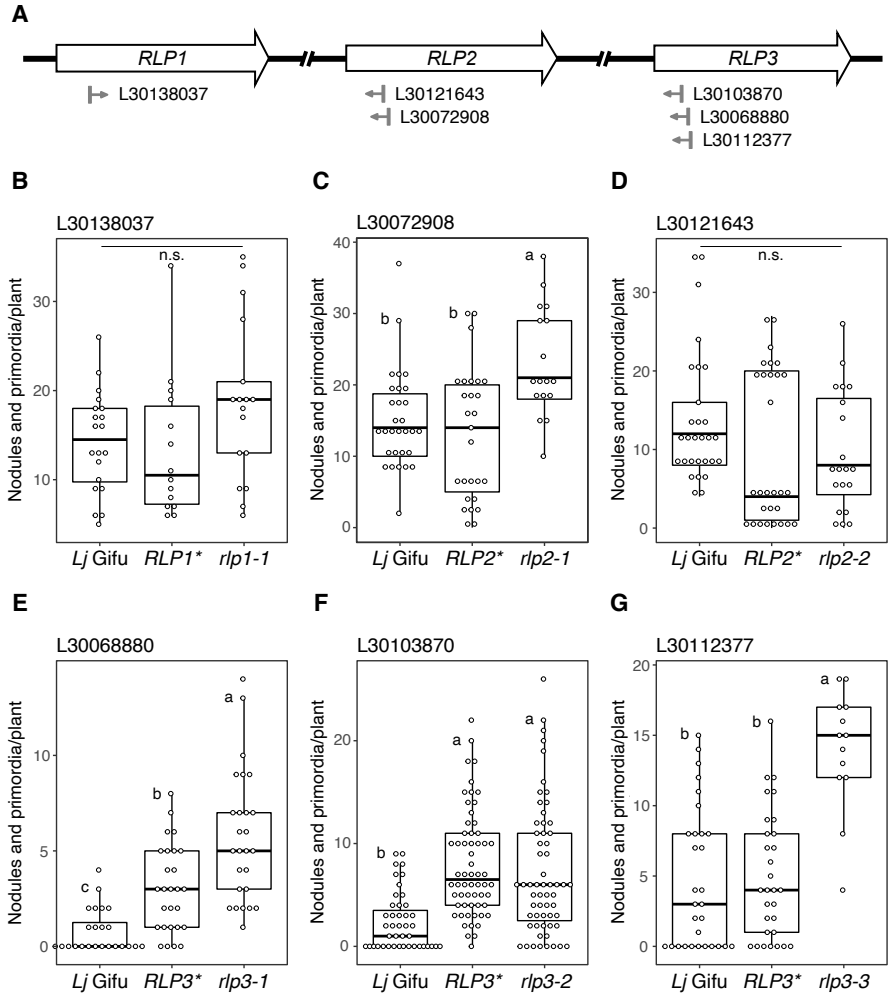
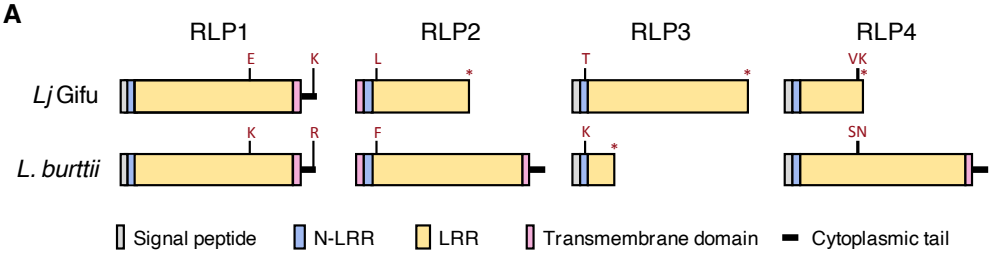


Figure 30. Phenotypic analysis of *RLP* mutant lines in *L. japonicus* Gifu (*Lj* Gifu). *Lotus* Retrotransposon 1 (LORE1) mutant lines were ordered from Lotus Base (https://lotus.au.dk/lore1/order). Lines were screened for *RLP* homozygous mutant (*rlp*) and *RLP* wild-type (*RLP**) with background mutations. Plants were grown in HM and observed at 6 wpi with *Rl* Norway (OD_{600} = 0.005). (A) The relative position of the LORE1 insertions in *RLP* genes. The insertions are depicted in grey with arrows indicating the direction of the LORE1 insertion. (B) Phenotyping results of *RLP1* mutant line, L30138037. (C) and (D) Phenotyping results of *RLP2* mutant lines, L30072908 and L30121643, respectively. (E) to (G) Phenotyping result of *rlp3*-mutant lines, L30068880, L30103870, and L30112377, respectively. Each dot represents one individual plant. The black bars in the box plot indicate the median of each condition. Different lowercase letters indicate significant differences determined by ANOVA followed by Tukey's test (p <0.05) (n.s., not significant).

A



B

Transformation	TM	*oxRLP2*	*oxRLP4*[b]	*oxRLP2+4*[a]	*npRLP2*[b]	*npRLP4*	*npRLP2+4*
				Lj Gifu			
Nodulated/Total plants	16/60	12/60	26/60	29/60	4/50	16/70	20/60
Nodulation rate	0.27	0.20	0.43	0.48	0.08	0.23	0.33
Nodule/primordia number	4/20	7/7	26/32	25/55	0/7	6/28	16/30
Nodulations/Nodulated plants	1.50	1.17	2.35	2.76	1.75	2.13	2.30
Nodulaitons/Total plants	0.40	0.23	1.02	1.33	0.14	0.9	0.7

Transformation	TM	*oxRLP2*	*oxRLP4*	*oxRLP2+4*	*npRLP2*	*npRLP4*	*npRLP2+4*[a]
				rlp3-2			
Nodulated/Total plants	13/60	5/30	7/50	16/50	9/60	13/70	22/60
Nodulation rate	0.22	0.17	0.14	0.32	0.15	0.19	0.37
Nodule/primordia number	3/17	2/4	5/14	21/28	6/14	0/15	22/43
Nodulations/Nodulated plants	1.54	1.00	2.71	3.06	2.22	1.15	2.95
Nodulations/Total plants	0.33	0.20	0.38	0.98	0.33	1.15	1.08

C

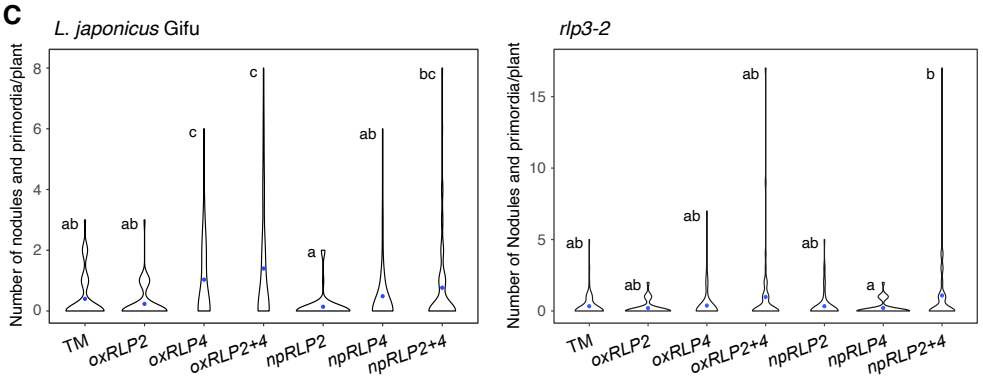
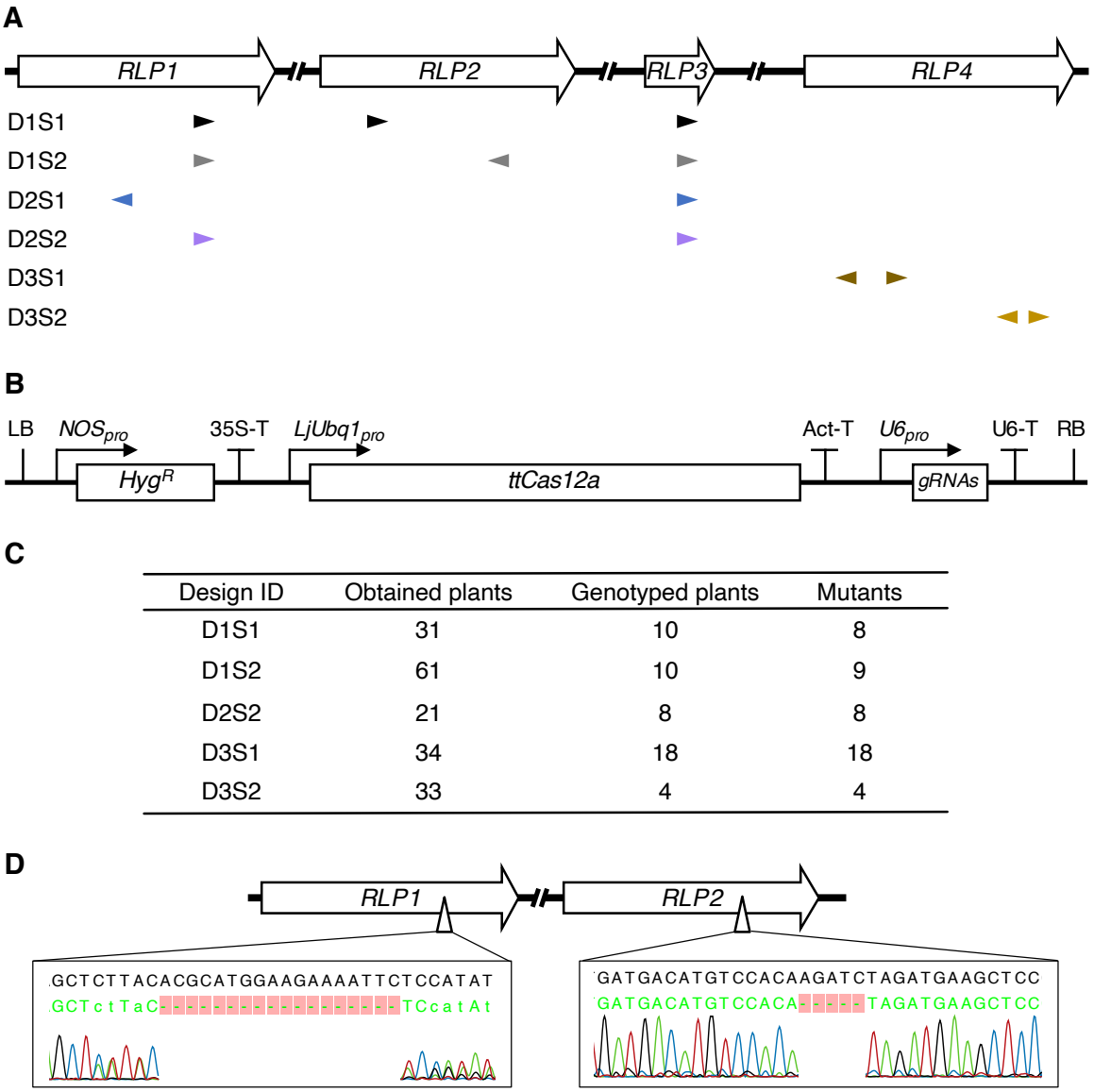


Figure 31. The effect of *LbRLP2* and *LbRLP4* trans-complementation on the nodulation of *L. japonicus* Gifu with *RI* Norway. (A) Predicted domain structure of candidate RLPs in *L. burttii* and *L. japonicus* Gifu (*Lj* Gifu). Both species have full-length RLP1 protein with 99.8 % sequence identity. Amino acid differences between RLPs of the different species are marked in red. Early stop codons are indicated by asterisks. (B) Results of trans-complementation assay. Each construct containing *RLP2* and/or *RLP4* driven by over-expression promoter ($AtUbq10_{pro}$ driven *LbRLP2* and $LjUbq1_{pro}$ driven *LbRLP4*) or native promoter ($LbRLP2_{pro}$ and $LbRLP4_{pro}$) was transiently transformed into *L. japonicus* Gifu or *rlp3-2*, L30103870. The nodulation rate was calculated by dividing nodulated plants over the total number of plants. Kruskal-Wallis test followed by Mann-Whitney test was used for statistical analysis on the nodulation rate of *RLP* trans-complementation with TM. Significant difference is labelled in lowercase letters (a, $p < 0.05$; b, $p < 0.005$). (C) The number of nodules and nodule primordia observed in trans-complemented plants. The transformed plants were observed at 8 wpi with *RI* Norway. Kruskal-Wallis test followed by Dunn's post-hoc-test was used for statistical analysis on numbers of nodules and primordia between conditions. Significant difference is labelled in lowercase letters ($p < 0.05$). TM, transformation marker; ox, over-expression promoter; np, native promoter.

5.4 Establishing *RLP* mutant lines in *L. burttii* using CRISPR-ttCas12a gene editing

LORE1-mutant lines were not available for all *RLP*s, and no double or triple mutant of the *RLP*s was available either. Most importantly, LORE1 lines were only available in *L. japonicus* Gifu, not in *L. burttii*. To study the loss of function of RLPs, *L. burttii* is the better system. Thus, to explore if the *RLPs* were functionally redundant, *RLP* mutant lines were generated in *L. burttii* background using the CRISPR-Cas12a gene-editing method. The Cas12a of *Lachnospiraceae* sp. (*La*Cas12a) in the CRISPR-Cas12a system targets the genomic position by guide RNA (gRNA), cleaving DNA at sites distal to protospacer-adjacent motif (PAM) and introducing mutations in the genes. The discovery of the temperature-tolerant Cas12a (ttCas12a) improves the efficiency in mutation induction under high temperatures (Schindele et al., 2020).

A total of six CRISPR-ttCas12a constructs targeting the *RLP* genes were designed and cloned using the Golden Gate method (Binder et al., 2014). The six CRISPR-ttCas12a constructs were categorized into three designs based on the targeted gene(s). Each design comprised two different sets of CRISPR-ttCas12a constructs containing two or three gRNAs targeting the *RLP*s at different positions (Figure 32A). To create a mutation in *RLP1*, *RLP2*, and *RLP3*, the constructs of design 1 contained three gRNAs targeting the three *RLP*s, respectively (Figure 32A, design 1). The constructs of design 2 were aimed to delete the region between *RLP1* and *RLP3*, which comprised gRNAs targeting *RLP1* and *RLP3* (Figure 32A, design 2). The aim of design 3 was to mutate *RLP4*, which included two gRNAs targeting *RLP4* (Figure 32A, design 3). Each ttCas12a construct contained a cassette expressing gRNAs, a cassette expressing ttCas12a, and a cassette expressing hygromycin resistant gene for selection (Figure 32B). The ttCas12a constructs were transformed into hypocotyls of *L. burttii* to generate stable *rlp*-mutant lines. The hypocotyl explants were cultured in different nutrient media to regenerate mutant plants through stages of callus induction, shoot and root formation, elongation, and propagation in the greenhouse, progressively. We obtained a total of 180 plants for 5 sets of gRNA designs (Figure 32C). No plants were regenerated from design 2 set 1. Deletions in *RLP* genes were identified after sequencing. Among the 50 genotyped plants, 46 of them were homozygous or heterozygous mutants in *RLP*s. Examples of the acquired mutations are shown in Figure 32D. The genotyping is currently ongoing.



C

Design ID	Obtained plants	Genotyped plants	Mutants
D1S1	31	10	8
D1S2	61	10	9
D2S2	21	8	8
D3S1	34	18	18
D3S2	33	4	4

Figure 32. Generation of *RLP*-mutants in *L. burttii* by the CRISPR-ttCas12a gene-editing system. (A) Relative position of each guide RNA (gRNA) sequence to the *RLPs*. The arrowheads indicate the position of the gRNAs and their direction relative to the *RLP*s. Five different combinations of the gRNAs are depicted in different colors. D, design; S, set. (B) The ttCas12a and gRNA expression construct transformed into *L. burttii*. LB, left border; RB, right border; NOS_{pro}, nopaline synthase promoter; $LjUbq1_{pro}$, *L. japonicus* Ubiquitin promoter; $U6_{pro}$, *L. japonicus* U6 promoter; 35S-T, Cauliflower mosaic virus 35S terminator; Act-T, *Nicotiana benthamiana* Actin terminator; U6-T, *L. japonicus* U6 terminator; Hyg^R, hygromycin resistance gene; *Cas12a*, *ttCas12a* of *Lachnospiraceae bacterium* ND2006; gRNAs, guide RNA sequences. (C) The number of obtained plants from each design. The number of genotyped plants and obtained mutants are indicated in the table. (D) Deletions are found in *RLP1* and *RLP2* after sequencing. Genomic DNA was extracted from one leaf of regenerated plants after CRISPR-ttCas12a gene editing. PCR was used to amplify the gRNA targeting region. The PCR product was sequenced by Sanger sequencing. The arrowheads indicate the position where mutation happened.

Discussion

1 *Lotus* possess different levels of promiscuity in symbiont selection

L. japonicus Gifu and MG-20 are model legumes due to their small genomes (470Mb), short generation period, small plant size, and abundant seed production. They can be easily transformed by *Agrobacterium* and regenerated from tissue (Hashiguchi et al., 2011, Handberg & Stougaard, 1992). *L. japonicus* MG-20, Funakura, and *L. filicaulis* are three common parents for inter-accession crosses with *L. japonicus* Gifu. The fourth crossing partner of *L. japonicus* Gifu is *L. burttii* (Kawaguchi et al., 2005, Sandal et al., 2012). Unlike the genome of *L. japonicus* MG-20, which harbors a translocation event between chromosomes 1 and 2, the genome of *L. burttii* and *L. japonicus* Gifu share a similar chromosome morphology (Kawaguchi et al., 2005). Although *L. burttii* and *L. japonicus* Gifu are genetically close, *L. burttii* is more promiscuous in symbiont selection (Zarrabian et al., 2021). Zarrabian et al. described that *L. burttii* nodulated with 40 out of 42 tested rhizobia strains regardless if they led to a nitrogen-fixing interaction (Zarrabian et al., 2021). In this study, *L. burttii* also nodulated with the four tested strains regardless of whether a nitrogen-fixing interaction can be established and showed a higher number of nodules and nodule primordia with *Rl* Norway among the tested *Lotus* accessions (Figure 4, 5). Hence, we hypothesized *L. burttii* is less selective, as this increases its chances to meet its compatible symbionts.

Much of the current understanding about symbiotic partner selection is derived from the perspective of rhizobia, as they are relatively easy to be genetically manipulated (Walker et al., 2020). One clear difference that differentiates *L. burttii* from *L. japonicus* ecotypes is its compatibility with *Ef* HH103 (Figure 4), which is consistent with previous studies (Sandal et al., 2012, Acosta-Jurado et al., 2019, Acosta-Jurado et al., 2020). On the one hand, the *NodD2*, *NolR*, and *SyrM* of *Ef* HH103 have been identified that contribute to its compatibility with *L. burttii* and its incompatibility with *L. japonicus* Gifu. *Ef* HH103 with mutations in *NodD2*, *NolR*, or *SyrM* can form effective nodules on *L. japonicus* Gifu (Acosta-Jurado et al., 2019, Acosta-Jurado et al., 2020). On the other hand, *NFR1* was proposed to be a candidate that contributes to the determination of symbiotic compatibility, as two missense substitutions were found between *L. burttii* and *L. japonicus* Gifu (A124T and D213Y, *L. burttii*/*L. japonicus* Gifu) (Zarrabian et al., 2021). Introducing the extracellular domain of *LbNFR1* into wild-type *L. japonicus* Gifu and *Ljnfr1-1* mutant restored the *Ljnfr1-1* nodulation with *M. loti* R7A. However, this fails to extend the symbiotic compatibility of *L. japonicus* Gifu and did not result

in the *Ef* HH103 nodulation on *L. japonicus* Gifu plants in their experimental condition (Zarrabian et al., 2021). In this study, *Rl* Norway nodulated *L. burttii* and nodulated *L. japonicus* Gifu in HM but not LM (Figure 6 and 7). These results indicate that the genetic difference in *NFR1* between *L. burttii* and *L. japonicus* Gifu is not the only determinant of their difference in symbiotic compatibility. Altogether, symbiotic compatibility is likely to be determined by rhizobial genes, *Lotus* gene(s), and substrate moisture.

The promiscuity in symbiotic compatibility has been documented in different legume species and some of the genetic and molecular causes have been identified (Walker et al., 2020). *Rl* bv. *trifolii* strains that establish effective symbiosis with *T. subterraneum* can be ineffective or less effective on *T. repens* and *T. pretense* (Kumar et al., 2015, Tesfaye, 1998). Some *Rl* bv. *trifolii* strains that are compatible with *T. resupinatum* are not compatible with *T. rubens*, which is not strict in the symbiont selection and hosts rhizobia close to *R. grahamii* and *R. galegae* (Marek-Kozaczuk et al., 2017). In *G. max*, the presence of *Rj4* restricts the nodulation by many *B. japonicum* and *B. elkanii* strains (Vest & Caldwell, 1972, Sadowsky & Cregan, 1992, Tang et al., 2016). In the symbiosis between *Medicago* and *Sinorhizobium* strains, *S. meliloti* 1021 nodulates *M. sativa* and effectively fixes nitrogen. However, its nitrogen-fixing activity decreases in *M. truncatula* A17 and R108 (Terpolilli et al., 2008, Kazmierczak et al., 2017) and it fails to induce nodules in *M. murex* (Terpolilli et al., 2008). Instead, *M. truncatula* A17 and R108 establish efficient nitrogen-fixing symbiosis with *S. medicae* WSM419 and *S. meliloti* 102F34, respectively. A role of rhizobial EPS in the symbiotic compatibility between *Medicago* and *Sinorhizobium* strains has been found. Mutant strains of *S. meliloti* Rm1021 that present defective or altered production in EPS II fails to infect or ineffectively infect its host alfalfa (Cheng & Walker, 1998). In this study, the perception of EPS may not block the nodulation of *Rl* Norway in *Lotus* accessions, as they possess the ability to nodulate with *Rl* Norway either in LM condition or HM condition. However, different substrate moisture may affect EPS synthesis, which further influences the EPS perception and causes conditionally nodulation of *Rl* Norway with *L. japonicus* Gifu.

The perception of rhizobial Nod factors and defense response of hosts also play roles in symbiont selection. Amino acid changes in the extracellular LysM domain 1 of NFR1 and LysM domain 2 of NFR5 affect the recognition of the chitin core of Nod factors, which leads to the loss of Nod factor-induced signaling or the alteration of the rhizobial symbiosis partner (Radutoiu et al., 2003, Radutoiu et al., 2007, Bozsoki et al., 2020). The presence of *nod*, *noe*, and *nol*, which mediate the substitutions of Nod factors can either broaden or restrict the symbiotic compatibility (Firmin et al., 1993, Berck et al., 1999, D'Haeze & Holsters, 2002). A restriction of nodulation is found in Afghanistan peas, which can be overcome by *R.*

leguminosarum bv. *viciae* strain TOM that harbors the *nodX* gene (Firmin et al., 1993), whereas an extension on the host range of *E. fredii* USDA257 can be found by introducing the *nolL* of *Rhizobium* sp. strain NGR234 (Berck et al., 1999). As *L. burttii* nodulates with rhizobial strains producing Nod factors with a broad variety of decorations, it is less likely that the minor changes in Nod factor receptors lead to the striking difference in promiscuity between *L. burttii* and *L. japonicus* Gifu (Zarrabian et al., 2021).

The presence of resistant genes contributes to the incompatibility of symbionts (Vest & Caldwell, 1972, Yang et al., 2010, Zhang et al., 2021). While the presence of *Rj4* in soybean limits the nodulation by many *B. japonicum* and *B. elkanii* strains (Vest & Caldwell, 1972, Sadowsky & Cregan, 1992, Tang et al., 2016), the presence of *Rfg1* and *Rj2* blocks the nodulation by *Ef* USDA257, *Ef* USDA193, and *B. japonicum* USDA 122, respectively (Yang et al., 2010, Fan et al., 2017). Similarly, soybean cultivars comprising *Nodule Number Locus 1* (*GmNNL1*) cannot be nodulated by *B. diazoefficiens* USDA110 (Zhang et al., 2021). This symbiotic incompatibility results from the effector-triggered immunity of the host, which is triggered by rhizobial effectors, such as nodulation outer protein P (NopP) (Yang et al., 2010, Tsukui et al., 2013, Tsurumaru et al., 2015, Zhang et al., 2021, Sugawara et al., 2018). However, missing resistant genes may not be the cause for the promiscuity of *L. burttii*, since the F1 progeny of a cross between *L. burttii* and *L. japonicus* Gifu nodulated with *Rl*. Norway. These results suggest the role of additional factor(s) in selecting the symbionts.

2 Substrate moisture alters the symbiotic compatibility of *L. japonicus* Gifu with *Rl* Norway

2.1 Effect of different growth substrates on nodulation of *L. japonicus* Gifu

Plant-growth substrates have different physical and chemical properties, such as cation exchange capacity, pH buffering ability, and water absorption, retention and releasing capability (Asaduzzaman et al., 2015). Together with water availability, they are the most critical factors for plant growth and development (Jankauskienė et al., 2015, Yang et al., 2018). The effect of different growth substrates has been observed in plant growth, nutrient loss, and crop yield (Asaduzzaman et al., 2015, Jankauskienė et al., 2015, Yang et al., 2018). In this study, we inspected the effect of substrate and moisture on the *Lotus*-rhizobia symbiosis. Different substrates and HM environment altered the symbiotic compatibility of *L. japonicus* Gifu with *Rl* Norway from non-nodulation to nodulation (Figure 6, 7).

The characteristics of the different substrates may explain our observation. Clay granules, sand, and vermiculite are non-organic growth substrates with variations in particle size (Verdonck et al., 1982). Clay granules are either irregular or kidney-shaped with 4-8 mm in diameter, and vermiculite are 3-6 mm in size. Sand particles are relatively small around 1.4-2.2 mm in size. The differences between the tested substrates lead to a different degree of the volume-weight number (g/L) and total pore space (%), which are negatively correlated. Substrates that have higher volume-weight numbers have lower total pore spaces (Verdonck et al., 1982). Compared to the sand-vermiculite mixture, sand has the highest volume-weight number and the lowest total pore space while the clay granules have the opposite characteristics. As the pore space is positively correlated to the air permeability and negatively correlated to the moisture of the micro-environment, the most compact and loose pore spaces may be unfavorable for *L. japonicus* Gifu to nodulate with *Rl* Norway regardless of the amount of added growth media. The intermediate level of total pore space possessed by sand-vermiculite mixture probably provides the flexibility to adjust the micro-environment by adding different amounts of growth media. Thus, the nodulation and non-nodulation conditions of *L. japonicus* Gifu with *Rl* Norway could be distinguished. The effect of growth substrate has also been identified in the symbiosis between grassland plants and arbuscular mycorrhizal fungi (AMF). Decreasing the proportion of Haplic Chernozem soil and increasing the proportion of sand in the growth substrate promotes AMF colonization, leading to increased plant biomass. This phenomenon may result from the reduced nutrient limitation by AMF symbiosis in growth substrates containing more sand (Zaller et al., 2011). Altogether, these findings indicate that the substrate environment may have a dramatic impact on plant-microbe symbiosis.

2.2 Substrate moisture affects nodulation by mediating rhizobia colonization

In this study, we first observed the nodulation ability of *L. japonicus* Gifu with *Rl* Norway. Although *L. japonicus* Gifu showed minor nodulation with *Rl* Norway in different growth substrates in LM, its symbiotic compatibility significantly increased in sand-vermiculite under HM conditions. The increasing nodule numbers were coupled with the rising substrate moisture (Figure 7) and an additive effect between high substrate moisture and *Rl* Norway inoculation on nodulation was observed in spontaneous nodulation mutant lines, *snf1-1* and *snf2-2* (Figure 8). As rhizobia colonization is a prerequisite to trigger nodulation (Rivilla et al., 2017), we inspected if *Rl* Norway colonized the roots of *L. japonicus* Gifu differently between LM and HM conditions and discovered a positive correlation between substrate moisture, *Rl* Norway colonization, and *L. japonicus* Gifu nodulation. In LM condition, *Rl* Norway colonization was barely observed and *L. japonicus* Gifu did not nodulate. In HM condition, *L. japonicus* Gifu nodulated with *Rl* Norway and higher *Rl* Norway colonization was observed (Figure 9). As LM

was unfavorable and HM was more appropriate for *Rl* Norway colonization, it is possible that either a blocking event in LM or a promoting event in HM was elicited in *L. japonicus* Gifu, since *Rl* Norway nodulates *L. burttii* in LM.

Substrate moisture also greatly affects bacterial motility (Harshey, 2003) and sufficient water content in the substrate is crucial for bacterial root adhesion (Rodriguez-Navarro et al., 2007). Insufficient water content can limit the flagellar motility of bacteria and thus limit the bacteria colonization (Dechesne et al., 2010), which may explain the weak *Rl* Norway colonization in LM. The imbalanced osmotic pressure, raised solute concentration in the remaining water, and the limited diffusion of substrates due to LM may also impact bacteria bioactivities adversely and indirectly (Csonka, 1989, Schimel et al., 2007, Naylor & Coleman-Derr, 2017). Sand-vermiculite was used as the growth substrate in our experiments. As vermiculite is very porous, its strong capillary action and water-holding ability (Papadopoulos et al., 2008) could lead to fragmented aquatic microhabitats in LM (Or et al., 2007). These fragmented aquatic microhabitats are connected by thin liquid films that limit bacterial motility (Or et al., 2007). Therefore, bacteria with limited motility may be retained on the vermiculite particles and could not reach the roots of host plants in LM, which causes a lack of bacterial colonization and further leads to the absence of nodulation. The absence of *L. japonicus* Gifu nodulation in LM may also be a consequence of a defect in rhizobial infection, which results from the abnormal rhizobial EPS production, as the EPS production is highly dependent on soil moisture content and changes under drought conditions (Breedveld et al., 1990, Mendrygal & Gonzalez, 2000). As rhizobial EPS is required for early infection process (Fraysse et al., 2003, Laus et al., 2006), defected or altered EPS production can lead to deficient host infection (Rolfe et al., 1996, Cheng & Walker, 1998). Hence, limited bacterial motility or deficient host infection in combination with the potential blocking event elicited by host lead to the absence of nodulation on *L. japonicus* Gifu in LM, whereas the absence of host blocking event and adequate moisture for *Rl* Norway colonization lead to the nodulation of *L. japonicus* Gifu in HM.

2.3 Substrate moisture-mediated changes in metabolism at the transcriptomic level between nodulated and non-nodulated *L. japonicus* Gifu

The major transcriptional difference caused by substrate moisture between *L. japonicus* Gifu grown in LM and HM after *Rl* Norway inoculation was in genes related to the oxidation-reduction processes (Figure 12C, 13). Several DEGs were categorized in the GO terms related to oxidation-reduction processes and ROS metabolism. Oxidation-reduction processes are involved in cellular metabolism and plant adaptation to environmental changes and regulated by metabolic processes in return (Geigenberger & Fernie, 2014, Ribeiro et al., 2015). During

symbiosis establishment, host and rhizobia communicate with each other by mediating the host-produced ROS, the signaling molecules of redox processes. (Ribeiro et al., 2015). As the stage after rhizobial colonization and before nodule formation was examined in this study, we speculated that the predominant redox response in the transcriptome is neither due to rhizobia infection nor nodule senescence (Puppo et al., 2005, Damiani et al., 2016). Since redox processes are known to engage in cell division and organ formation (Mhamdi & Van Breusegem, 2018), the observed DEGs related to redox processes may be one cause that is responsible for the different nodulation phenotype between *L. japonicus* grown in LM and HM after *Rl* Norway inoculation.

ROS act as signal molecules that respond to abiotic stresses. Manipulating ROS levels can enhance abiotic stress tolerance (You & Chan, 2015). However, ROS become toxic to plants when their accumulation exceeds the amount of scavenging (Huang et al., 2019). In this study, we found the expression level of several genes encoding peroxidase, glutathione S-transferase, aldehyde dehydrogenase, MsrB, and ubiquinol oxidase were relatively high in LM condition (Supplementary File 2, 3). These enzymes are involved in redox homeostasis and antioxidation by either scavenging ROS or reducing oxidative stress (Kotchoni et al., 2006, Hasanuzzaman et al., 2020, Rey & Tarrago, 2018). The higher expression of these enzyme-encoding genes may correlate to higher antioxidant activity, suggesting that *Rl* Norway-inoculated *L. japonicus* Gifu may encounter oxidative stress or have higher ROS production in LM. As ROS and peroxidases also take part in cell wall lignification, cell wall stiffening and lossening (Liu, 2012, Somssich, 2020), the higher ROS levels and peroxidase expressions could lead to alteration in the cell wall structure, which may be responsible for deficient rhizobia colonization. In HM, the expression level of several genes encoding enzymes that affect cell wall structure and stability were relatively higher, such as genes encoding laccase, subtilisin-like protease, and pectin esterase (Supplementary File 2, 3). While subtilisin-like proteases influence the cell wall by regulating pectin esterase activity, laccases polymerize the monolignols, which become lignin polymers in the plant cell wall (Schaller et al., 2018, Berthet et al., 2012). This finding may be a hint to the nodulation phenotype in HM, as cell wall modification is required for nodule formation. Altogether, our findings suggest that *Rl* Norway triggered different metabolic responses at the transcriptomic level in *L. japonicus* Gifu in LM and HM.

The genes involved in carbohydrate-, amide-, peptide-and ion-related biosynthesis were differentially regulated by substrate moisture (Figure 12, 13), which may lead to changes in the production of primary and secondary metabolites. These changes may be coupled with the alteration in oxidation-reduction processes and ROS, as redox and ROS are greatly involved

in cellular metabolism and catabolism (Geigenberger & Fernie, 2014, Ribeiro et al., 2015). The changes in metabolite production could further lead to changes in transportation, as transporters are required to transport the metabolites, ions, and phytohormones (Theodoulou & Kerr, 2015). Similar to a transcriptomic study in peanuts, the GO terms related to redox, metabolic process, and catalytic activity are enriched in peanuts after bradyrhizobia infection and before nodule formation (Peng et al., 2017). This finding suggests the alteration in redox activity and metabolism is a shared feature in the different symbiotic systems at the stage after rhizobial infection and before nodule formation.

Carbohydrates, amides, and peptides are often primary and secondary metabolites in root exudates (Bais et al., 2006, Badri & Vivanco, 2009). They serve as a source of nutrients for rhizospheric microbes (Compant et al., 2010) and attract beneficial rhizobacteria for plants to adapt to abiotic stresses (Bobille et al., 2019, Feng et al., 2021). The root exudates can also influence the interplay between plants and their rhizospheric microbes via physical, chemical, or biological interactions (Haichar et al., 2014). The composition and quantity of the root exudates vary depending on the plant species, age, and biotic and abiotic conditions (Haichar et al., 2008, Lombardi et al., 2018, Cesari et al., 2019). Water deficiency is one of the abiotic factors that can alter the profile of root exudates (Calvo et al., 2017). When the water supply is reduced, the number of carbohydrates, amino acids, potassium, electric conductivity, and hormone concentrations in the root exudates are increased (Calvo et al., 2017, Bobille et al., 2019, Hartman et al., 2019). An alteration of the root exudate composition may be elicited by substrate moisture, which then affects the symbiotic compatibility of *L. japonicus* Gifu with *Rl* Norway in LM and HM. Analysis of the root metabolome and root exudates is an important future research direction (in collaboration with J. Ranner and C. Dawid, Technical University of Munich).

2.4 Substrate moisture-mediated transcriptomic changes in phytohormone-related genes

Phytohormones play essential roles for plants to adapt to the changing environment. ABA, ethylene, JA, cytokinins, and GAs have been reported to mediate the plant defense response against pathogens and abiotic stress (Verma et al., 2016, Wani et al., 2016). Plants accumulate ethylene and inhibit ABA biosynthesis in response to waterlogging (Wang et al., 2021), whereas ABA is one of the main regulators of water deficiency (Finkelstein, 2013). Here we found that few genes involved in ethylene, JA, cytokinin, and GA biosynthesis were categorized in the oxidative-reduction processes GO category, but no ABA-related genes. While ethylene and JA biosynthesis genes were preferentially expressed in LM rather than HM,

genes related to GA and cytokinin were preferentially expressed in HM (Figure 14). Although it is difficult to define a drought condition, the molecular evidence suggests that the water content in the LM condition is not sufficient for *L. japonicus* Gifu, thus resulting in inhibition on nodulation with *Rl* Norway.

2.4.1 Ethylene

Ethylene is a volatile phytohormone regulating symbiosis establishment at different levels, including infection thread formation, total nodule number, nodule morphology, and nodule positioning (Ferguson & Mathesius, 2014). The effect of ethylene on nodulation has been broadly documented in several legume species, comprising *Pisum*, *Medicago*, *Lotus,* and *Trifolium* (Goodlass & Smith, 1979, Nukui et al., 2000, Oldroyd et al., 2001, Lee & Larue, 1992). Since previous literature indicates that ethylene negatively regulates the nodulation and we found *ACO*, the gene encoding the rate-limiting enzyme of ethylene biosynthesis (Houben & Van de Poel, 2019) is preferentially expressed in LM (Figure 14), ethylene becomes our candidate that may be responsible for the non-nodulation of *L. japonicus* Gifu with *Rl* Norway in HM. To investigate the effect of ethylene on nodulation, we applied the ethylene precursor ACC or the ethylene biosynthesis inhibitor AVG to *L. japonicus* Gifu plants inoculated with *Rl* Norway and phenotyped their nodulation. No significant difference between plants with either treatment was found in our experimental setup (Figure 15), even though ACC treatment has been reported to inhibit nodule formation and Ca^{2+} spiking in *M. truncatula* root hairs in response to NF (Oldroyd, 2001). ACC is often a replacement of gaseous ethylene for laboratory convenience, but the decomposition of ACC to ethylene is slow and the ethylene releasing rate is experimentally uncontrollable. In addition, ACC can be consumed in a short period after application (Zhang et al., 2010, Zhang & Wen, 2010). More importantly, the gene encoding ACO that converts ACC into ethylene was downregulated in HM (Figure 14). Therefore, ACC treatment may not be the ideal replacement of ethylene for long-term and quantitative experiments (Tucker & Wen, 2015). Improvement in experimental design would be required to study the role of ethylene in the nodulation of *L. japonicus* Gifu.

2.4.2 JA, GAs and cytokinin

JA, together with ABA and ethylene, is involved in osmotic adjustment and drought-related responses (Ullah et al., 2018). GAs comprise a large group of tetracyclic diterpenoid carboxylic acids, of which GA_1 to GA_4 are bioactive (Sponsel & Hedden, 2010). JA accumulation and GA reduction can enhance drought tolerance (Bandurska et al., 2003, Sanchez-Romera et al., 2014, Li et al., 2012). Still, the effect of GA application remains debatable. It either induces (Kawaguchi et al., 1996) or inhibits (Maekawa et al., 2009) the nodulation in *L. japonicus.* In

our transcriptomic data, the JA biosynthesis gene *OPR* exhibited higher expression in LM, and the *GA3OX2*, which converts GA precursors into the bioactive GAs, showed lower expression in LM (Figure 14). This finding implies a higher JA and a lower GA level in LM, which is consistent with the literature. Cytokinin is also involved in nodule development. A cytokinin dehydrogenase-encoding gene, *CKX3*, is induced by the Nod factors at the early stage of nodule initiation. This enzyme functions in cytokinin degradation and is responsible for cytokinin homeostasis (Reid et al., 2016). Reduced nodulation is found in *ckx3* mutants, which is caused by the elevated cytokinin (Reid et al., 2016). We found this gene was preferentially expressed in HM (Figure14), implying higher cytokinin degradation activity in HM than in LM. As a higher endogenous cytokinin level inhibits nodulation, this finding suggested that the HM condition is more favorable than LM for *L. japonicus* Gifu to nodulate with *Rl* Norway. However, whether changing the expression of one gene can affect the whole phytohormone production is obscure.

2.5 Flavonoids enhance the symbiotic compatibility of *L. japonicus* Gifu with *Rl* Norway in high substrate moisture

Flavonoids are secondary metabolites broadly distributed in plants. They act as antioxidants (Williams et al., 2004) and are involved in plant defense response against biotic and abiotic stresses (Winkel-Shirley, 2001, Bidart-Bouzat & Imeh-Nathaniel, 2008). Flavonoids are known to induce rhizobia *nod* genes expression and Nod factor production (Zhang et al., 2009, Abdel-Lateif et al., 2012). Enzymes encoded by *CHS*, *F3H*, *FLS*, and *PKR* are part of the flavonoid biosynthesis pathway. These genes, especially *CHS*, were downregulated in HM in our transcriptomes (Figure 16). This finding is contradictory to the previous study in *Medicago*. Zhang et al. found that silencing *CHS* of *M. truncatula* leads to a great reduction in its nodulation ability with *S. meliloti* due to its flavonoid-deficient roots (Zhang et al., 2009). In contrast, we observed a downregulation of *CHS* in the nodulation condition. This contradiction may result from the different research systems as we inspected the sub-compatible symbiosis whereas they examined the compatible symbiosis.

The downregulation of flavonoid biosynthesis genes may lead to the decrease of flavonoid end products and the accumulation of intermediates. Thus, a higher level of kaempferol and quercetin in LM and a higher level of *p*-coumaroyl-CoA and naringenin in HM are suggested. As kaempferol and quercetin accumulate under drought stress (Nakabayashi et al., 2014) and their higher content is a characteristic for drought tolerance (Sarker & Oba, 2020), the water content in our LM condition may be an unfavorable condition for *L. japonicus* Gifu to nodulate. Although the role of *p*-coumaroyl-CoA in symbiosis is limited, it can be converted into lignin

that contributes to cell wall lignification and abiotic resistance (Dong & Lin, 2021). Higher lignin production may suggest that the LM condition is stressful to *L. japonicus* Gifu and altering the level of cell wall lignification may result in an unfavorable root condition for *Rl* Norway colonization. Naringenin is an inducer of *nod* genes of *Rhizobium* spp., such as *nodA, nodC*, and *nodD* of *R. leguminosarum* strains (Zaat et al., 1987, Begum et al., 2001) and *nodSU* of *Rhizobium* sp. NGR234 (Lewin et al., 1990). A growth-promoting effect of naringenin has been reported in *Bradyrhizobium* sp. strain ORS285 *in vitro*, and this growth-promoting effect is suggested to facilitate rhizobial colonization (Nouwen et al., 2019). Thus, we inspected the role of naringenin in the nodulation of *L. japonicus* Gifu with *Rl* Norway. In line with the literature, naringenin induced the $nodA_{pro}$ of *Rl* Norway (Figure 18, 19). The colonization-promoting effect of naringenin was observed (Figure 20), but neither the growth-promoting effect on *Rl* Norway nor the effect on nodulation was shown (Figure 17). Hence, naringenin may promote the nodulation of *L. japonicus* Gifu by increasing the colonization of *Rl* Norway in HM rather than promoting the nodulation directly. However, one limitation should be noted. In this study, the naringenin was applied in the growth substrate at the beginning of the *Rl* Norway inoculation without supplementation until the phenotyping point at 6 wpi. Whether the amount of naringenin in the enclosed culture system is enough to affect the nodulation is debatable. Modification in experimental design may provide a better view of the effect of naringenin on nodulation.

3 RLP-encoding genes contribute to the different symbiotic compatibility in *Lotus*

3.1 *Rl* Norway induces a symbiotic response in *L. burttii* and not a defense response in *L. japonicus* Gifu

Our previous results demonstrated that HM promotes the *Rl* Norway colonization and nodulation in *L. japonicus* Gifu. However, the transcriptome of *L. japonicus* Gifu plants in HM does not resemble the transcriptome of *L. burttii* plants (Figure 11), which suggests a host-dependent factor. Our transcriptomic analysis indicated that *L. burttii* activated the symbiotic response to *Rl* Norway, while *Rl* Norway generally induced changes in biological processes of *L. japonicus* Gifu in LM (Figure 22-24). Symbiotic marker genes, such as *NPL*, *NIN*, *ERN1*, and *EPR3*, and genes involved in cell wall metabolic processes specifically upregulated in response to rhizobia were upregulated in *L. burttii* (Figure 22, 24). These findings are in line with a previous study in peanut. Peanut plants that nodulate with *Bradyrhizobium* sp. exhibits higher *NIN* and *ERN1* expression, whereas their sister inbred lines that cannot nodulate with *Bradyrhizobium* sp. do not activate the symbiosis signaling genes (Peng et al., 2017). A

previous study demonstrates that *Rl* Norway can induce GUS activity in the *L. japonicus* Gifu reporter line T90, which responds to rhizobia and Nod factor (Gossmann et al., 2012). Although this is contradictory to our finding, the difference between the two studies should be noted. Here we inspected the gene expression in transcriptome globally, whereas a symbiosis-specific promoter was inspected in the previous study. Our transcriptomes also revealed several genes encoding senescence-related proteins were upregulated in *L. japonicus* Gifu (Figure 22). Although previous studies demonstrate the senescence regulation in nodule senescence (Zhou et al., 2021), the exact function of those senescence-related proteins remains unclear. Hence, whether the senescence response is potently involved in symbiotic compatibility determination and represses the nodulation of *L. japonicus* Gifu in LM is obscure. Overall, *Rl* Norway induced symbiotic response in *L. burttii* but not in *L. japonicus* Gifu in LM.

Suppression of the rhizobia-elicited host defense responses is required for symbiosis establishment (Gourion et al., 2015). Two inducible host defense systems are microbe-associated (or pathogen-associated) molecular pattern (MAMP/PAMP)-triggered immunity (MTI/PTI) and effector-triggered immunity (ETI) (Nishad et al., 2020). The perception of MAMPs by host pattern-recognition receptors activates MTI at the early stage of symbiotic interaction (Gourion et al., 2015). While previous research reports the rhizobial Nod factor-mediated MTI suppression in non-legume species (Liang et al., 2013), many studies show that effectors are released to counteract MTI (Nishad et al., 2020). However, effectors can also elicit ETI. In rhizobia, effectors can be delivered to the host via type III (T3SS), type IV (T4SS), or type VI (T6SS) secretion systems (Gourion et al., 2015, Safronova et al., 2020). Compared to MTI, ETI is stronger and often leads to hypersensitive cell death (Gourion et al., 2015, Nishad et al., 2020). Host resistance proteins can either directly recognize the effectors via a receptor-ligand binding mechanism or indirectly sense effector-mediated modifications of endogenous proteins or decoys (Cui et al., 2015).

The interaction of effectors and ETI influences symbiotic compatibility, which has been studied in different legume-rhizobia symbioses. For instance, soybean with the presence of *Rj2* inhibits the nodulation by *B. japonicum* USDA122 (Caldwell, 1966, Caldwell et al., 1966), which is resulted from the NopP effector-elicited immune response (Sugawara et al., 2018). This ETI-mediated restriction also applies to *Rj4* and *GmNNL1* of soybean, which limits the nodulation by many *B. japonicum*, *B. elkanii* strains, and *B. diazoefficiens* USDA110 (Vest & Caldwell, 1972, Sadowsky & Cregan, 1992, Zhang et al., 2021). In *Lotus*, a T3SS-mutant strain of *B. elkanii* USDA61 fails to infect *L. japonicus* Gifu and has defects in nodule maturation on *L. burttii* (Kusakabe et al., 2020), while the *Ef* HH103 T3SS-mutant obtains the compatibility to nodulate *L. japonicus* Gifu (Jimenez-Guerrero et al., 2020). T4SS is suggested to be an analog

of T3SS regarding its role in symbiosis and the exclusivity between the presence of T3SS and T4SS is found, in which some *Lotus* symbionts that possess T3SS do not have T4SS, and vice versa (Hubber et al., 2004, Paco et al., 2019). T6SS has been identified in *R. leguminosarum*, *M. loti*, and *E. fredii* and is suggested to affect the host specificity, however, this remains understudied (Bingle et al., 2008, Sugawara et al., 2013, Safronova et al., 2020). Since *Rl* Norway harbors T4SS and T6SS, it can potentially release effectors that are perceived by *L. japonicus* Gifu and block the nodulation. However, *L. japonicus* Gifu inoculated with *Rl* Norway strains defective in T4SS and T6SS remain non-nodulated (S. Masachis, personal communication). In addition, our transcriptomic data did not support this hypothesis either. Most of the inspected defense marker genes showed similar and/or low expression levels in both *L. burttii* and *L. japonicus* Gifu (Figure 24). Moreover, *Rl* Norway successfully nodulated the F1 progeny of *L. burttii* and *L. japonicus* Gifu cross (Figure 25), indicating the dominancy of the nodulation phenotype of *L. burttii*. Hence, it is unlikely that *L. japonicus* Gifu represses the *Rl* Norway nodulation due to the defense response.

3.2 RLPs contribute to the difference in symbiotic compatibility between *L. burttii* and *L. japonicus* Gifu

To reveal the genes that contribute to the symbiotic compatibility, a QTL mapping was performed on RILs between *L. burttii* and *L. japonicus* Gifu. An 0.3 Mb region was correlated with the different nodulation phenotypes (Zarrabian et al., 2021). Our microsynteny analysis suggested that this QTL is syntenic between *Lotus*, *Medicago*, and *Phaseolus* (Figure 26). The same region in *Medicago* is syntenic to the *Sym2* region in *Pisum* (Limpens et al., 2003, Gualtieri et al., 2002). Within the *Sym2* region, a LysM-RLK encoding gene *LykX* was identified that contributes to the determination of symbiotic compatibility of Afghanistan with *R. leguminosarum* bv. *viciae* (Solovev et al., 2021). These findings suggest that the QTL region is conserved in legumes species and may also contribute to the different symbiotic compatibility in *Lotus*.

Within the QTL, three LysM-RLK encoding genes were previously identified (in collaboration with S.U. Anderson). Since preliminary trans-complementation assays of the *L. burttii LysM-RLK* orthologs into *L. japonicus* Gifu did not confer nodulation (S.U. Anderson and Fan-Yu Yu, personal communication), the existence of other candidate genes is implied. As the nodulation phenotype of *L. burttii* is dominant, we hypothesized that the candidate genes comprise a full-length CDS in *L. burttii.* By SNP identification, five candidate genes that comprised a full-length CDS in *L. burttii* were identified. However, two of them encoded uncharacterized proteins and one gene encoded a ribonuclease H, which was located in a highly repetitive genome region.

Based on previous studies that indicate the role of LRR-RLP-encoding genes in plant-microbe interaction (Jones et al., 1994, Gust & Felix, 2014, Jamieson et al., 2018) and our examination on the expression pattern in transcriptome and by qRT-PCR, two LRR-RLP encoding genes, *LbRLP2* and *LbRLP4*, were considered as primary candidates (Figure 27-29). The LRR-RLPs are a major group of RLPs characterized by their extracellular leucine-rich repeats (LRR) as a ligand-binding domain. In addition, they comprise a transmembrane domain and a short cytoplasmic tail (Jamieson et al., 2018). Most of the reported LRR-RLPs function in plant immunity, growth and development while current knowledge of LRR-RLPs in symbiosis remains limited (Gust & Felix, 2014, Jamieson et al., 2018). For example, *Cf-9* of *Lycopersicon esculentum* is the first-identified resistance gene encoding an LRR-RLP that recognizes the Avr9 protein produced by some races of fungus *Cladosporium fulvum* (Jones, 1994). In *Arabidopsis*, TOO MANY MOUTHS (TMM) balance stem cell renewal and differentiation to regulate stomatal distribution (Nadeau & Sack, 2002). In this study, we showed that the trans-complementation of *LbRLP4* and *LbRLP2+4* into *L. japonicus* Gifu increases its symbiotic compatibility with *Rl* Norway in LM (Figure 31). While the *Lb*RLP4 is the major contributor, the presence of *Lb*RLP2 elevates the LbRLP4-mediated symbiotic compatibility in LM. Moreover, the *RLP1* is highly conserved whereas the SNPs in *RLP3* result in a dramatic length difference in CDS between *L. burttii* and *L. japonicus* Gifu. Hence, we proposed that *LbRLP4* is required for the symbiotic compatibility with *Rl* Norway.

Due to the lack of kinase domain, RLPs often associate with RLKs to initiate or attenuate the signal transduction (Jamieson et al., 2018). The SUPPRESSOR OF BIR1-1/ EVERSHED (SOBIR1/EVR) is an example of LRR-RLK that works as an adaptor kinase for plant LRR-RLPs. It associates with multiple tomato RLPs (Cf proteins) and is involved in defense response against fungal pathogens (Liebrand et al., 2013). To activate the RLP/adaptor complexes, a ligand-binding step with co-receptors like BAK1 or SERKs is required (Gust & Felix, 2014). Often LRR type proteins and kinases associate with each other. Yet, the Nod factor perceiving complex formed by LysM-type receptors NFR1 and NFR5, associates with the LRR-type RLK named symbiosis receptor-like kinase (SYMRK) for downstream signaling transduction (Stracke et al., 2002). Thus, we postulated an additional RLP-RLK complex, which contributes to the different symbiotic compatibility between *L. burttii* and *L. japonicus* Gifu (Figure 33). The Nod factor perception triggers downstream signaling pathways in *L. burttii* and *L. japonicus* Gifu. However, additional signals may be required for *Rl* Norway nodulation in the two *Lotus* species. An additional RLP-RLK complex comprising RLP1, RLP2, RLP4, and an unidentified RLK may perceive a putative molecule of *Rl* Norway, which provides an additional signal for *L. burttii* to nodulate with *Rl* Norway. In *L. japonicus* Gifu, RLP2 and RLP4 are truncated, which fail to form an RLP-RLK complex that transduces the downstream

signal. Both *Lb*RLP3 and *Lj*RLP3 are without transmembrane domain but the length of *Lj*RLP3 is five times longer than *Lb*RLP3. We hypothesize that *Lj*RLP3 may recognize the putative *Rl* Norway molecule and keep it away from the recognition RLP-RLK complex. Without the additional signal from the RLP-RLK complex, *L. japonicus* Gifu is unable to nodulate with *Rl* Norway in LM. Under HM conditions, an additional high moisture-dependent signal is provided. Thus, *L. japonicus* Gifu nodulates with *Rl* Norway in HM.

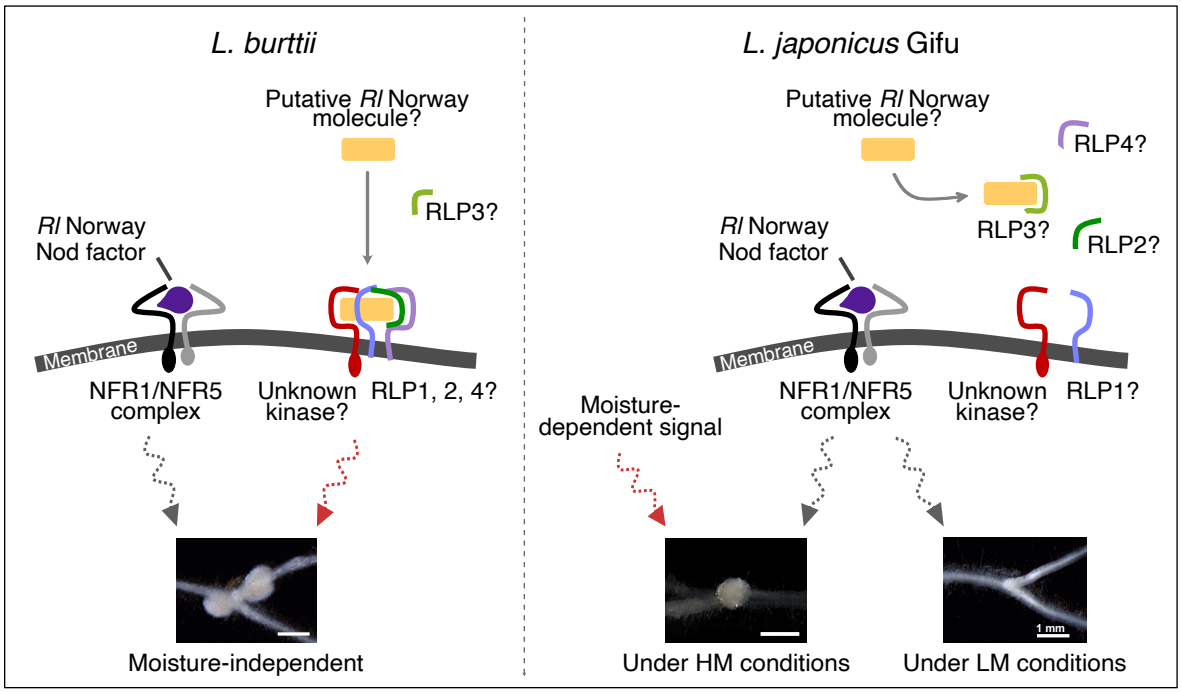


Figure 33. Proposed regulatory mechanism of the different symbiotic compatibility in *L. burttii* and *L. japonicus* Gifu with *Rl* Norway. In addition to the Nod factor-induced signals, additional signals are proposed to be required for *Rl* Norway nodulation in *L. burttii* and *L. japonicus* Gifu, which are induced by a putative *Rl* Norway molecule and by high substrate moisture, respectively. In *L. burttii*, a putative *Rl* Norway molecule is recognized by *Lb*RLP1, *Lb*RLP2, and *Lb*RLP4 complex together with an unknown kinase to induce downstream signal transduction for the nodulation with *Rl* Norway. In *L. japonicus* Gifu, early stop codons are found in *LjRLP2* and *LjRLP4*, which may encode truncated RLP proteins and make the RLP-RLK complex fail to generate the signals. Moreover, *Lj*RLP3 may recognize the putative *Rl* Norway molecule and keep it away from the recognition RLP-RLK complex as its length is five times longer than *Lb*RLP3, although both of them are without the transmembrane domain. Due to the lack of additional signals, *Rl* Norway is unable to nodulate *L. japonicus* Gifu in low substrate moisture. However, additional signals ascribing to higher moisture in the growth environment contribute to enhancing the compatibility. Hence, *Rl* Norway can nodulate *L. japonicus* Gifu when environmental moisture is above a certain threshold. LM, low substrate moisture; HM, high substrate moisture.

Conclusion

A study in 2012 demonstrated that *Rl* Norway infects and nodulates a broad range of *Lotus* species (Gossmann et al., 2012). In this study, we demonstrated various symbiotic compatibility between *Lotus* accessions and four sub-compatible rhizobia strains and studied the genetic and abiotic determinants that contribute to the variation of symbiotic compatibility between *Lotus* accessions and its sub-compatible symbiont *Rl* Norway. Although previous studies indicate *Rl* Norway is unable to nodulate *L. japonicus* Gifu (Gossmann et al., 2012), we found that HM alters the symbiotic compatibility of *L. japonicus* Gifu with *Rl* Norway from non-nodulation to nodulation. We proposed a high moisture-dependent signal from *L. japonicus* Gifu grown in HM is required for its nodulation with *Rl* Norway (Figure 33). With transcriptomic analysis and phenotypic observations, we pointed out the potential regulatory mechanisms in *L. japonicus* Gifu and suggested that the difference in the composition of root exudates between LM and HM is the cause of nodulation suppression in LM. This provides an enticing future research direction.

L. burttii and *L. japonicus* Gifu show a striking difference in nodulation with *Rl* Norway when grown in LM. The transcriptomic data demonstrate the symbiotic markers and cell wall mediating genes were upregulated regulated in *L. burttii*, whereas senescence-related genes were upregulated in *L. japonicus* Gifu. Further genetic studies indicate the nodulation phenotype of *L. burttii* is dominant and center an 0.3 Mb QTL correlated with nodulation phenotype. We improved the genome annotation in the QTL and identified *LbRLP2* and *LbRLP4* as the most promising candidates leading to the variation of nodulation phenotype. By trans-complementing *LbRLP2* and *LbRLP4* into *L. japonicus* Gifu, the symbiotic compatibility of *L. japonicus* Gifu with *Rl* Norway increased. We proposed an RLP-RLK complex that recognizes a putative *Rl* Norway molecule and provides an additional signal for nodulation in addition to NFR1/NFR5 Nod factor perception complex (Figure 33). Moreover, mutants in the *RLP* cluster of *L. burttii* were generated by the CRISPR-Cas12a gene-editing system. These materials provide opportunities to validate the effect of *RLP*s in the future.

References

Abdel-Lateif K, Bogusz D, Hocher V, 2012. The role of flavonoids in the establishment of plant roots endosymbioses with arbuscular mycorrhiza fungi, rhizobia and *Frankia* bacteria. *Plant Signal Behav* **7**, 636-41.

Acosta-Jurado S, Alias-Villegas C, Navarro-Gomez P*, et al.*, 2020. *Sinorhizobium fredii* HH103 *syrM* inactivation affects the expression of a large number of genes, impairs nodulation with soybean and extends the host-range to *Lotus japonicus*. *Environ Microbiol* **22**, 1104-24.

Acosta-Jurado S, Fuentes-Romero F, Ruiz-Sainz JE, Janczarek M, Vinardell JM, 2021. Rhizobial exopolysaccharides: genetic regulation of their synthesis and relevance in symbiosis with legumes. *Int J Mol Sci* **22**.

Acosta-Jurado S, Rodriguez-Navarro DN, Kawaharada Y*, et al.*, 2019. *Sinorhizobium fredii* HH103 *nolR* and *nodD2* mutants gain capacity for infection thread invasion of *Lotus japonicus* Gifu and *Lotus burttii*. *Environ Microbiol* **21**, 1718-39.

Ahmed M, Rauf M, Akhtar M, Mukhtar Z, Saeed NA, 2020. Hazards of nitrogen fertilizers and ways to reduce nitrate accumulation in crop plants. *Environ Sci Pollut Res Int* **27**, 17661-70.

Akashi T, Sawada Y, Shimada N, Sakurai N, Aoki T, Ayabe S-I, 2003. cDNA cloning and biochemical characterization of S-adenosyl-L-methionine: 2, 7, 4′-trihydroxyisoflavanone 4′-O-methyltransferase, a critical enzyme of the legume isoflavonoid phytoalexin pathway. *Plant Cell Physiol.* **44**, 103-12.

Alexa A, Rahnenfuhrer J, 2019. *topGO: Enrichment Analysis for Gene Ontology*. R package version 2.36.0.

Allen EK, Allen ON, 1950. Biochemical and symbiotic properties of the rhizobia. *Bacteriol Rev* **14**, 273-330.

Amor BB, Shaw SL, Oldroyd GE*, et al.*, 2003. The *NFP* locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. *Plant J* **34**, 495-506.

Antolin-Llovera M, Ried MK, Parniske M, 2014. Cleavage of the symbiosis receptor-like kinase ectodomain promotes complex formation with Nod factor receptor 5. *Curr Biol* **24**, 422-7.

Ariel F, Brault-Hernandez M, Laffont C*, et al.*, 2012. Two direct targets of cytokinin signaling regulate symbiotic nodulation in *Medicago truncatula*. *Plant Cell* **24**, 3838-52.

Arora N, Skoog F, Allen ON, 1959. Kinetin-induced pseudonodules on tobacco roots. *Am J Bot* **46**, 610-3.

Arrese-Igor C, Royuela M, Lorenzo C, Felipe MR, Aparicio-Tejo PM, 1993. Effect of low rhizosphere oxygen on growth, nitrogen fixation and nodule morphology in lucerne. *Physiol Plant* **89**, 55-63.

Arrighi JF, Barre A, Ben Amor B*, et al.*, 2006. The *Medicago truncatula* lysin motif-receptor-like kinase gene family includes *NFP* and new nodule-expressed genes. *Plant Physiol* **142**, 265-79.

Arthikala MK, Montiel J, Sanchez-Lopez R, Nava N, Cardenas L, Quinto C, 2017. Respiratory burst oxidase homolog gene a is crucial for rhizobium infection and nodule maturation and function in common bean. *Front Plant Sci* **8**, 2003.

Asaduzzaman M, Saifullah M, Mollick AKMSR, Hossain MM, Halim GMA, Asao T, 2015. Influence of soilless culture substrate on improvement of yield and produce quality of horticultural crops. In *Soilless Culture – Use of Substrates for the Production of Quality Horticultural Crops.*

Badri DV, Vivanco JM, 2009. Regulation and function of root exudates. *Plant Cell Environ* **32**, 666-81.

Bagnoli JW, Ziegenhain C, Janjic A*, et al.*, 2018. Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat Commun* **9**, 2937.

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM, 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* **57**, 233-66.

Ballizany WL, Hofmann RW, Jahufer MZZ, Barrett BA, 2012. Multivariate associations of flavonoid and biomass accumulation in white clover (*Trifolium repens*) under drought. *Funct Plant Biol* **39**, 167-77.

Bandurska H, Stroiński A, Kubiś J, 2003. The effect of jasmonic acid on the accumulation of ABA, proline and spermidine and its influence on membrane injury under water deficit in two barley genotypes. *Acta Physiol Plant* **25**, 279-85.

Bano A, Harper JE, 2002. Plant growth regulators and phloem exudates modulate root nodulation of soybean. *Funct Plant Biol* **29**, 1299-307.

Bao Y, Aggarwal P, Robbins NE, 2nd*, et al.*, 2014. Plant roots use a patterning mechanism to position lateral root branches toward available water. *Proc Natl Acad Sci U.S.A.* **111**, 9319-24.

Beeckman F, Motte H, Beeckman T, 2018. Nitrification in agricultural soils: impact, actors and mitigation. *Curr Opin Biotechnol* **50**, 166-73.

Begum AA, Leibovitch S, Migner P, Zhang F, 2001. Specific flavonoids induced nod gene expression and pre-activated nod genes of *Rhizobium leguminosarum* increased pea (*Pisum sativum* L.) and lentil (*Lens culinaris* L.) nodulation in controlled growth chamber environments. *J Exp Bot* **52**, 1537-43.

Bellato C, Krishnan HB, Cubo T, Temprano F, Pueppke SG, 1997. The soybean cultivar specificity gene *nolX* is present, expressed in a *nodD*-dependent manner, and of symbiotic significance in cultivar-nonspecific strains of *Rhizobium* (*Sinorhizobium*) *fredii*. *Microbiology* **143**, 1381-8.

Beltran JP, Canas LA, 2018. Grain and forage legumes: nutritional value and agriculture sustainability. *Methods Mol Biol* **1822**, 1-10.

Berck S, Perret X, Quesada-Vincens D, Prome J, Broughton WJ, Jabbouri S, 1999. *NolL* of *Rhizobium* sp. strain NGR234 is required for O-acetyltransferase activity. *J Bacteriol* **181**, 957-64.

Beringer JE, 1974. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* **84**, 188-98.

Bertani G, 1951. Studies on lysogenesis. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol* **62**, 293-300.

Berthet S, Thevenin J, Baratiny D*, et al.*, 2012. Chapter 5 – Role of plant laccases in lignin polymerization. In *Lignins – Biosynthesis, Biodegradation and Bioengineering* **61,** 145-72.

Betti M, Perez-Delgado C, Garcia-Calderon M, Diaz P, Monza J, Marquez AJ, 2012. Cellular stress following water deprivation in the model legume *Lotus japonicus*. *Cells* **1**, 1089-106.

Bhattacharyya A, Pablo CHD, Mavrodi OV, Weller DM, Thomashow LS, Mavrodi DV, 2021. Rhizosphere plant-microbe interactions under water stress. *Adv Appl Microbiol* **115**, 65-113.

Bhattacharyya PN, Jha DK, 2012. Plant growth-promoting rhizobacteria (PGPR): emergence in agriculture. *World J Microbiol Biotechnol* **28**, 1327-50.

Bidart-Bouzat MG, Imeh-Nathaniel A, 2008. Global change effects on plant chemical defenses against insect herbivores. *J Integr Plant Biol* **50**, 1339-54.

Binder A, Lambert J, Morbitzer R*, et al.*, 2014. A modular plasmid assembly kit for multigene expression, gene silencing and silencing rescue in plants. *PLoS One* **9**, e88218.

Bingle LE, Bailey CM, Pallen MJ, 2008. Type VI secretion: a beginner's guide. *Curr Opin Microbiol* **11**, 3-8.

Bobille H, Fustec J, Robins RJ, Cukier C, Limami AM, 2019. Effect of water availability on changes in root amino acids and associated rhizosphere on root exudation of amino acids in *Pisum sativum* L. *Phytochemistry* **161**, 75-85.

Bordeleau LM, Prévost D, 1994. Nodulation and nitrogen fixation in extreme environments. *Plant Soil* **161**, 115-25.

Borisov AY, Madsen LH, Tsyganov VE*, et al.*, 2003. The *Sym35* gene required for root nodule development in pea is an ortholog of *NIN* from *Lotus japonicus*. *Plant Physiol* **131**, 1009-17.

Borowik A, Wyszkowska J, 2016. Soil moisture as a factor affecting the microbiological and biochemical activity of soil. *Plant Soil Environ* **62**, 250-5.

Bottomley PJ, Myrold DD, 2007. Biological N Inputs. In *Soil Microbiology, Ecology and Biochemistry.* 365-87.

Bozsoki Z, Gysel K, Hansen SB*, et al.*, 2020. Ligand-recognizing motifs in plant LysM receptors are major determinants of specificity. *Science* **369**, 663-70.

Breedveld MW, Zevenhuizen LPTM, Zehnder AJB, 1990. Osmotically induced oligo- and polysaccharide synthesis by *Rhizobium meliloti* SU-47. *J Gen Microbiol* **136**, 2511-9.

Bruning B, Rozema J, 2013. Symbiotic nitrogen fixation in legumes: Perspectives for saline agriculture. *Environmental and Experimental Botany* **92**, 134-43.

Caldwell BE, 1966. Inheritance of a strain-specific ineffective nodulation in soybeans. *Crop Sci* **6**, 427-8.

Caldwell BE, Hinson K, Johnson HW, 1966. A strain-specific ineffective nodulation reaction in the soybean *Glycine max* L. Merrill. *Crop Sci* **6**, 495-6.

Calvo OC, Franzaring J, Schmid I, Muller M, Brohon N, Fangmeier A, 2017. Atmospheric CO_2 enrichment and drought stress modify root exudation of barley. *Glob Chang Biol* **23**, 1292-304.

Capoen W, Goormachtig S, Holsters M, 2010. Water-tolerant legume nodulation. *J Exp Bot* **61**, 1251-5.

Cardenas L, Martinez A, Sanchez F, Quinto C, 2008. Fast, transient and specific intracellular ROS changes in living root hair cells responding to Nod factors (NFs). *Plant J* **56**, 802-13.

Cerri MR, Wang Q, Stolz P*, et al.*, 2017. The *ERN1* transcription factor gene is a target of the CCaMK/CYCLOPS complex and controls rhizobial infection in *Lotus japonicus*. *New Phytol* **215**, 323-37.

Cesari A, Paulucci N, Lopez-Gomez M, Hidalgo-Castellanos J, Pla CL, Dardanelli MS, 2019. Restrictive water condition modifies the root exudates composition during peanut-PGPR interaction and conditions early events, reversing the negative effects on plant growth. *Plant Physiol Biochem* **142**, 519-27.

Charpentier M, 2018. Calcium signals in the plant nucleus: origin and function. *J Exp Bot* **69**, 17.

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. *Plant Cell* **20**, 3467-79.

Charpentier M, Sun J, Vaz Martins T*, et al.*, 2016. Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science* **352**, 1102-5.

Chen S, Perathoner S, Ampelli C, Centi G, 2019. Chapter 2 – Electrochemical Dinitrogen Activation: to find a sustainable way to produce ammonia. In *Studies in Surface Science and Catalysis* **178**, 31-46.

Cheng HP, Walker GC, 1998. Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. *J Bacteriol* **180**, 5183-91.

Cho MJ, Harper JE, 1993. Effect of abscisic acid application on root isoflavonoid concentration and nodulation of wild-type and nodulation-mutant soybean plants. *Plant Soil* **153**, 145-9.

Choudhury SR, Pandey S, 2015. Phosphorylation-dependent regulation of G-protein cycle during nodule formation in soybean. *Plant Cell* **27**, 3260-76.

Coba De La Peña T, Pueyo JJ, 2011. Legumes in the reclamation of marginal soils, from cultivar and inoculant selection to transgenic approaches. *Agron Sustain Dev* **32**, 65-91.

Compant S, Clément C, Sessitsch A, 2010. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. *Soil Biol Biochem* **42**, 669-78.

Concordet JP, Haeussler M, 2018. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* **46**, W242-W5.

Cooper JE, Wood M, Bjourson AJ, 1985. Nodulation of *Lotus pedunculatus* in acid rooting solution by fast- and slow-growing rhizobia. *Soil Biol Biochem* **17**, 487-92.

Csonka LN, 1989. Physiological and genetic responses of bacteria to osmotic stress. *Microbiol Rev* **53**, 121-47.

Cui H, Tsuda K, Parker JE, 2015. Effector-triggered immunity: from pathogen perception to robust defense. *Annu Rev Plant Biol* **66**, 487-511.

D'haeze W, Holsters M, 2002. Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology* **12**, 79R-105R.

Daher FB, Braybrook SA, 2015. How to let go: pectin and plant cell adhesion. *Front Plant Sci* **6**, 523.

Damiani I, Pauly N, Puppo A, Brouquisse R, Boscari A, 2016. Reactive oxygen species and nitric oxide control early steps of the legume – rhizobium symbiotic interaction. *Front Plant Sci* **7**, 454.

De Lajudie P, Laurent-Fulele E, Willems A*, et al.*, 1998. *Allorhizobium undicola* gen. nov., sp. nov., nitrogen-fixing bacteria that efficiently nodulate *Neptunia natans* in Senegal. *Int J Syst Bacteriol* **48**, 1277-90.

De Lyra Mdo C, Lopez-Baena FJ, Madinabeitia N*, et al.*, 2006. Inactivation of the *Sinorhizobium fredii* HH103 *rhcJ* gene abolishes nodulation outer proteins (Nops) secretion and decreases the symbiotic capacity with soybean. *Int Microbiol* **9**, 125-33.

Debelle F, Plazanet C, Roche P*, et al.*, 1996. The NodA proteins of *Rhizobium meliloti* and *Rhizobium tropici* specify the N-acylation of Nod factors by different fatty acids. *Mol Microbiol* **22**, 303-14.

Dechesne A, Wang G, Gulez G, Or D, Smets BF, 2010. Hydration-controlled bacterial motility and dispersal on surfaces. *Proc Natl Acad Sci U.S.A.* **107**, 14369-72.

Dechorgnat J, Nguyen CT, Armengaud P*, et al.*, 2010. From the soil to the seeds: the long journey of nitrate in plants. *J Exp Bot* **62**, 1349-59.

Del Castillo LD, Hunt S, Layzell DB, 1994. The role of oxygen in the regulation of nitrogenase activity in drought-stressed soybean nodules. *Plant Physiol* **106**, 949-55.

Dimkpa CO, Fugice J, Singh U, Lewis TD, 2020. Development of fertilizers for enhanced nitrogen use efficiency – Trends and perspectives. *Sci Total Environ* **731**, 139113.

Ding Y, Kalo P, Yendrek C*, et al.*, 2008. Abscisic acid coordinates nod factor and cytokinin signaling during the regulation of nodulation in *Medicago truncatula*. *Plant Cell* **20**, 2681-95.

Dobin A, Davis CA, Schlesinger F*, et al.*, 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21.

Dolgikh AV, Kirienko AN, Tikhonovich IA, Foo E, Dolgikh EA, 2019. The DELLA proteins influence the expression of cytokinin biosynthesis and response genes during nodulation. *Front Plant Sci* **10**, 432.

Dong NQ, Lin HX, 2021. Contribution of phenylpropanoid metabolism to plant development and plant-environment interactions. *J Integr Plant Biol* **63**, 180-209.

Downie JA, 1998. Functions of rhizobial nodulation genes. In *The Rhizobiaceae: molecular biology of model plant-associated bacteria.* 387-402.

El Msehli S, Rima N, Sghaier H, Aschi-Smiti S, Brouquisse R, 2016. Impact of hypoxia on nodulation and growth of the legume plant *Medicago truncatula* in symbiosis with *Sinorhizobium meliloti*. *Int J Curr Res Bios Plant Bio* **3**, 53-60.

Fan Y, Liu J, Lyu S, Wang Q, Yang S, Zhu H, 2017. The soybean *Rfg1* gene restricts nodulation by *Sinorhizobium fredii* USDA193. *Front Plant Sci* **8**, 1548.

Fang Y, Hirsch AM, 1998. Studying early nodulin gene *ENOD40* expression and induction by nodulation factor and cytokinin in transgenic alfalfa. *Plant Physiol* **116**, 53-68.

Farooq U, Bano A. Effect of abscisic acid and chlorocholine chloride on nodulation and biochemical content of *Vigna radiata* L. under water stress, *2007*.

Fedoroff NV, Battisti DS, Beachy RN*, et al.*, 2010. Radically rethinking agriculture for the 21st century. *Science* **327**, 833-4.

Feng H, Fu R, Hou X*, et al.*, 2021. Chemotaxis of beneficial rhizobacteria to root exudates: The first step towards root-microbe rhizosphere interactions. *Int J Mol Sci* **22**.

Ferguson BJ, Foo E, Ross JJ, Reid JB, 2011. Relationship between gibberellin, ethylene and nodulation in *Pisum sativum*. *New Phytol* **189**, 829-42.

Ferguson BJ, Mathesius U, 2014. Phytohormone regulation of legume-rhizobia interactions. *J Chem Ecol* **40**, 770-90.

Ferguson BJ, Mens C, Hastwell AH*, et al.*, 2019. Legume nodulation: The host controls the party. *Plant Cell Environ* **42**, 41-51.

Ferguson BJ, Ross JJ, Reid JB, 2005. Nodulation phenotypes of gibberellin and brassinosteroid mutants of pea. *Plant Physiol* **138**, 2396-405.

Fernandez-Gobel TF, Deanna R, Munoz NB, Robert G, Asurmendi S, Lascano R, 2019. Redox systemic signaling and induced tolerance responses during soybean-*Bradyrhizobium japonicum* interaction: involvement of Nod factor receptor and autoregulation of nodulation. *Front Plant Sci* **10**, 141.

Finkelstein R, 2013. Abscisic acid synthesis and response. *Arabidopsis Book* **11**, e0166.

Firmin JL, Wilson KE, Carlson RW, Davies AE, Downie JA, 1993. Resistance to nodulation of cv. Afghanistan peas is overcome by *nodX*, which mediates an O-acetylation of the *Rhizobium leguminosarum* lipo-oligosaccharide nodulation factor. *Mol Microbiol* **10**, 351-60.

Fliegmann J, Bono JJ, 2015. Lipo-chitooligosaccharidic nodulation factors and their perception by plant receptors. *Glycoconj J* **32**, 455-64.

Folch-Mallol JL, Manyani H, Marroqui S*, et al.*, 1998. Sulfation of nod factors via *nodHPQ* is *nodD* independent in *Rhizobium tropici* CIAT899. *Mol Plant Microbe Interact* **11**, 979-87.

Fonouni-Farde C, Diet A, Frugier F, 2016a. Root development and endosymbioses: DELLAs lead the orchestra. *Trends Plant Sci* **21**, 898-900.

Fonouni-Farde C, Tan S, Baudin M*, et al.*, 2016b. DELLA-mediated gibberellin signalling regulates Nod factor signalling and rhizobial infection. *Nat Commun* **7**, 12636.

Foo E, Plett JM, Lopez-Raez JA, Reid D, 2019. Editorial: The role of plant hormones in plant-microbe symbioses. *Front Plant Sci* **10**, 1391.

Foyer CH, Lam HM, Nguyen HT*, et al.*, 2016. Neglecting legumes has compromised human health and sustainable food production. *Nat Plants* **2**, 16112.

Foyer CH, Nguyen H, Lam HM, 2019. Legumes – The art and science of environmentally sustainable agriculture. *Plant Cell Environ* **42**, 1-5.

Fraysse N, Couderc F, Poinsot V, 2003. Surface polysaccharide involvement in establishing the rhizobium-legume symbiosis. *Eur J Biochem* **270**, 1365-80.

Fred EB, Baldwin IL, Mccoy E, 1932. Chapter 8 - Species relationships. In *Root nodule bacteria and leguminous plants*.

Freiberg C, Fellay R, Bairoch A, Broughton WJ, Rosenthal A, Perret X, 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* **387**, 394-401.

Friesen ML, 2012. Widespread fitness alignment in the legume-rhizobium symbiosis. *New Phytol* **194**, 1096-111.

Frink CR, Waggoner PE, Ausubel JH, 1999. Nitrogen fertilizer: retrospect and prospect. *Proc Natl Acad Sci U.S.A.* **96**, 1175-80.

Garcia-Calderon M, Perez-Delgado CM, Palove-Balang P, Betti M, Marquez AJ, 2020. Flavonoids and isoflavonoids biosynthesis in the model legume *Lotus japonicus*; connections to nitrogen metabolism and photorespiration. *Plants (Basel)* **9**, 774.

Garcia-Calderon M, Pons-Ferrer T, Mrazova A*, et al.*, 2015. Modulation of phenolic metabolism under stress conditions in a *Lotus japonicus* mutant lacking plastidic glutamine synthetase. *Front Plant Sci* **6**, 760.

Garwood EA, Ryden JC, 1986. Nitrate loss through leaching and surface runoff from grassland: effects of water supply, soil type and management. In *Developments in Plant and Soil Sciences* **23**.

Geigenberger P, Fernie AR, 2014. Metabolic control of redox and redox control of metabolism in plants. *Antioxid Redox Signal* **21**, 1389-421.

Geremia RA, Mergaert P, Geelen D, Van Montagu M, Holsters M, 1994. The NodC protein of *Azorhizobium caulinodans* is an N-acetylglucosaminyltransferase. *Proc Natl Acad Sci U.S.A.* **91**, 2669-73.

Geurts R, Heidstra R, Hadri AE*, et al.*, 1997. *Sym2* of pea is involved in a nodulation factor-perception mechanism that controls the infection process in the epidermis. *Plant Physiol* **115**, 351-9.

Gonzalez-Rizzo S, Crespi M, Frugier F, 2006. The *Medicago truncatula* CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with *Sinorhizobium meliloti*. *Plant Cell* **18**, 2680-93.

Goodlass G, Smith KA, 1979. Effects of ethylene on root extension and nodulation of pea (*Pisum sativum* L.) and white clover (*Trifolium repens* L.). *Plant Soil* **51**, 387-95.

Goormachtig S, Capoen W, James EK, Holsters M, 2004. Switch from intracellular to intercellular invasion during water stress-tolerant legume nodulation. *Proc Natl Acad Sci U.S.A.* **101**, 6303-8.

Gossmann JA, Markmann K, Brachmann A, Rose LE, Parniske M, 2012. Polymorphic infection and organogenesis patterns induced by a *Rhizobium leguminosarum* isolate from *Lotus* root nodules are determined by the host genotype. *New Phytol* **196**, 561-73.

Gourion B, Berrabah F, Ratet P, Stacey G, 2015. Rhizobium-legume symbioses: the crucial role of plant immunity. *Trends Plant Sci* **20**, 186-94.

Graff A, Conrad R, 2005. Impact of flooding on soil bacterial communities associated with poplar (*Populus* sp.) trees. *FEMS Microbiol Ecol* **53**, 401-15.

Granada CE, Strochein M, Vargas LK, Bruxel M, De Sa EL, Passaglia LM, 2014. Genetic diversity and symbiotic compatibility among rhizobial strains and *Desmodium incanum* and *Lotus* spp. plants. *Genet Mol Biol* **37**, 396-405.

Griffiths RI, Whiteley AS, O'donnell AG, Bailey MJ, 2003. Physiological and community responses of established grassland bacterial populations to water stress. *Appl Environ Microbiol* **69**, 6961-8.

Grobbelaar N, Clarke B, Hough MC, 1971. The nodulation and nitrogen fixation of isolated roots of *Phaseolus vulgaris* L. *Plant Soil* **35**, 215-23.

Groth M, Takeda N, Perry J*, 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 endosymbiotic development. *Plant Cell* **22**, 2509-26.

Grunewald W, Van Noorden G, Van Isterdael G, Beeckman T, Gheysen G, Mathesius U, 2009. Manipulation of auxin transport in plant roots during *Rhizobium* symbiosis and nematode parasitism. *Plant Cell* **21**, 2553-62.

Gualtieri G, Kulikova O, Limpens E*, et al.*, 2002. Microsynteny between pea and *Medicago truncatula* in the *SYM2* region. *Plant Mol Biol* **50**, 225-35.

Guinel FC, 2015. Ethylene, a hormone at the center-stage of nodulation. *Front Plant Sci* **6**, 1121.

Gust AA, Felix G, 2014. Receptor like proteins associate with SOBIR1-type of adaptors to form bimolecular receptor kinases. *Curr Opin Plant Biol* **21**, 104-11.

Haichar FEZ, Santaella C, Heulin T, Achouak W, 2014. Root exudates mediated interactions belowground. *Soil Biology and Biochemistry* **77**, 69-80.

Haichar FZ, Marol C, Berge O*, et al.*, 2008. Plant host habitat and root exudates shape soil bacterial community structure. *ISME J* **2**, 1221-30.

Hakeem KR, Sabir M, Ozturk M, Akhtar MS, Ibrahim FH, 2017. Nitrate and nitrogen oxides: sources, health effects and their remediation. *Rev Environ Contam Toxicol* **242**, 183-217.

Hanahan D, 1983. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-80.

Handberg K, Stougaard J, 1992. *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. *Plant J* **2**, 487-96.

Handberg, K., Stiller, J., Thykjær, T., & Stougaard, J. (1994). Transgenic Plants: *Agrobacterium*-mediated transformation of the diploid legume *Lotus japonicus*. *Cell Bio*, 119-127.

Harshey RM, 2003. Bacterial motility on a surface: many ways to a common goal. *Annu Rev Microbiol* **57**, 249-73.

Hartman K, Tringe SG, 2019. Interactions between plants and soil shaping the root microbiome under abiotic stress. *Biochem J* **476**, 2705-24.

Hartman S, Liu Z, Van Veen H*, et al.*, 2019. Ethylene-mediated nitric oxide depletion pre-adapts plants to hypoxia stress. *Nat Commun* **10**, 4020.

Hasanuzzaman M, Bhuyan M, Parvin K*, et al.*, 2020. Regulation of ROS metabolism in plants under environmental stress: A review of recent experimental evidence. *Int J Mol Sci* **21**.

Hashiguchi M, Tsuruta S-I, Akashi R, 2011. Morphological traits of *Lotus japonicus* (Regal) ecotypes collected in Japan. *Interdiscip Bio Cent* **3**, 4.1-4.7.

Hastwell AH, Gresshoff PM, Ferguson BJ, 2015. The structure and activity of nodulation-suppressing CLE peptide hormones of legumes. *Funct Plant Biol* **42**, 229-38.

Hayashi S, Gresshoff PM, Ferguson BJ, 2014. Mechanistic action of gibberellins in legume nodulation. *J Integr Plant Biol* **56**, 971-8.

He C, Gao H, Wang H*, et al.*, 2021. *GSK3*-mediated stress signaling inhibits legume-rhizobium symbiosis by phosphorylating *GmNSP1* in soybean. *Mol Plant* **14**, 488-502.

Heckmann AB, Lombardo F, Miwa H*, et al.*, 2006. *Lotus japonicus* nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume. *Plant Physiol* **142**, 1739-50.

Heckmann AB, Sandal N, Bek AS*, et al.*, 2011. Cytokinin induction of root nodule primordia in *Lotus japonicus* is regulated by a mechanism operating in the root cortex. *Mol Plant Microbe Interact* **24**, 1385-95.

Hedden P, Thomas SG, 2012. Gibberellin biosynthesis and its regulation. *Biochem J* **444**, 11-25.

Henry A, Doucette W, Norton J, Bugbee B, 2007. Changes in crested wheatgrass root exudation caused by flood, drought, and nutrient stress. *J Environ Qual* **36**, 904-12.

Hernández S, Rebuffo M, Arrivillaga S*, et al.*, 2005. Evaluation of the genotype-environment interaction in the establishment of *Lotus uliginosus* (Schkuhr) with soil-cores. *Lotus Newsletter* **35**, 120-30.

Herrera-Cervera J, 1999. At least five rhizobial species nodulate *Phaseolus vulgaris* in a Spanish soil. *FEMS Microbiol Ecol* **30**, 87-97.

Hirsch AM, Fang Y, Asad S, Kapulnik Y, 1997. The role of phytohormones in plant-microbe symbioses. *Plant Soil* **194**, 171-84.

Houben M, Van De Poel B, 2019. 1-Aminocyclopropane-1-Carboxylic Acid Oxidase (ACO): The enzyme that makes the plant hormone ethylene. *Front Plant Sci* **10**, 695.

Huang H, Ullah F, Zhou DX, Yi M, Zhao Y, 2019. Mechanisms of ROS regulation of plant development and stress responses. *Front Plant Sci* **10**, 800.

Hubber A, Vergunst AC, Sullivan JT, Hooykaas PJ, Ronson CW, 2004. Symbiotic phenotypes and translocated effector proteins of the *Mesorhizobium loti* strain R7A VirB/D4 type IV secretion system. *Mol Microbiol* **54**, 561-74.

Hunter S, Apweiler R, Attwood TK*, et al.*, 2009. InterPro: the integrative protein signature database. *Nucleic Acids Res* **37**, D211-5.

Ibanez F, Wall L, Fabra A, 2017. Starting points in plant-bacteria nitrogen-fixing symbioses: intercellular invasion of the roots. *J Exp Bot* **68**, 1905-18.

Imaizumi-Anraku H, Takeda N, Charpentier M*, et al.*, 2005. Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. *Nature* **433**, 527-31.

Insausti M, Timmis R, Kinnersley R, Rufino MC, 2020. Advances in sensing ammonia from agricultural sources. *Sci Total Environ* **706**, 135124.

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

Israel DW, Mathis JN, Barbour WM, Elkan GH, 1986. Symbiotic effectiveness and host-strain interactions of *Rhizobium fredii* USDA 191 on different soybean cultivars. *Appl Environ Microbiol* **51**, 898-903.

Jamieson PA, Shan L, He P, 2018. Plant cell surface molecular cypher: receptor-like proteins and their roles in immunity and development. *Plant Sci* **274**, 242-51.

Janczarek M, 2011. Environmental signals and regulatory pathways that influence exopolysaccharide production in rhizobia. *Int J Mol Sci* **12**, 7898-933.

Janczarek M, Rachwał K, Marzec A, Grządziel J, Palusińska-Szysz M, 2015. Signal molecules and cell-surface components involved in early stages of the legume-rhizobium interactions. *Appl Soil Ecol* **85**, 94-113.

Janjic A, Wange LE, Bagnoli JW*, et al.*, 2021. Prime-Seq, efficient and powerful bulk RNA-sequencing. DOI: 10.1101/2021.09.27.459575.

Jankauskienė J, Brazaitytė A, Viškelis P, 2015. Effect of different growing substrates on physiological processes, productivity and quality of tomato in soilless culture. In *Soilless Culture – Use of Substrates for the Production of Quality Horticultural Crops.*

Jarvis BDW, Pankhurst CE, Patel JJ, 1982. *Rhizobium loti*, a new species of legume root nodule bacteria. *Int J Syst Evol Microbiol* **32**, 378-80.

Jensen LS, Schjoerring JK, Van Der Hoek KW*, et al.*, 2011. Benefits of nitrogen for food, fibre and industrial production. In *The European Nitrogen Assessment: Sources, Effects and Policy Perspectives.* Cambridge University Press, 32-61.

Jimenez-Guerrero I, Acosta-Jurado S, Medina C*, et al.*, 2020. The *Sinorhizobium fredii* HH103 type III secretion system effector NopC blocks nodulation with *Lotus japonicus* Gifu. *J Exp Bot* **71**, 6043-56.

Jin Y, Liu H, Luo D*, et al.*, 2016. DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. *Nat Commun* **7**, 12433.

John M, Rohrig H, Schmidt J, Wieneke U, Schell J, 1993. *Rhizobium* NodB protein involved in nodulation signal synthesis is a chitooligosaccharide deacetylase. *Proc Natl Acad Sci U.S.A.* **90**, 625-9.

Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JD, 1994. Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* **266**, 789-93.

Ju C, Chang C, 2015. Mechanistic Insights in ethylene perception and signal transduction. *Plant Physiol* **169**, 85-95.

Kabdullayeva T, Crosbie DB, Marin M, 2020. *Mesorhizobium norvegicum* sp. nov., a rhizobium isolated from a *Lotus corniculatus* root nodule in Norway. *Int J Syst Evol Microbiol* **70**, 388-96.

Kalo P, Gleason C, Edwards A*, et al.*, 2005. Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science* **308**, 1786-9.

Kamal N, Mun T, Reid D*, et al.*, 2020. Insights into the evolution of symbiosis gene copy number and distribution from a chromosome-scale *Lotus japonicus* Gifu genome sequence. *DNA Res* **27**, 1-11.

Kamst E, Pilling J, Raamsdonk LM, Lugtenberg BJ, Spaink HP, 1997. Rhizobium nodulation protein NodC is an important determinant of chitin oligosaccharide chain length in Nod factor biosynthesis. *J Bacteriol* **179**, 2103-8.

Kanamori N, Madsen LH, Radutoiu S*, et al.*, 2006. A nucleoporin is required for induction of Ca^{2+} spiking in legume nodule development and essential for rhizobial and fungal symbiosis. *Proc Natl Acad Sci U.S.A.* **103**, 359-64.

Kaneko T, Nakamura Y, Sato S*, et al.*, 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti. *DNA Res* **7**, 331-8.

Kawaguchi M, Imaizumi-Anraku H, Fukai S, Syono K, 1996. Unusual branching in the seedlings of *Lotus japonicus* – Gibberellins reveal the nitrogen-sensitive cell divisions within the pericycle on roots. *Plant Cell Physiol* **37**, 461-70.

Kawaguchi M, Pedrosa-Harand A, Yano K*, et al.*, 2005. *Lotus burttii* takes a position of the third corner in the lotus molecular genetics triangle. *DNA Res* **12**, 69-77.

Kawaharada Y, Kelly S, Nielsen MW*, et al.*, 2015. Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature* **523**, 308-12.

Kazmierczak T, Nagymihaly M, Lamouche F*, et al.*, 2017. Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium* strains. *Mol Plant Microbe Interact* **30**, 399-409.

Kishorekumar R, Bulle M, Wany A, Gupta KJ, 2019. An overview of important enzymes involved in nitrogen assimilation of plants. *Methods Mol Biol* **2057**, 1-13.

Kotchoni SO, Gachomo EW, 2009. A rapid and hazardous reagent free protocol for genomic DNA extraction suitable for genetic studies in plants. *Mol Biol Rep* **36**, 1633-6.

Kotchoni SO, Kuhns C, Ditzer A, Kirch HH, Bartels D, 2006. Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to abiotic stress and protects plants against lipid peroxidation and oxidative stress. *Plant Cell Environ* **29**, 1033-48.

Krishnan HB, Oehrle NW, Alaswad AA*, et al.*, 2019. Biochemical and anatomical investigation of *Sesbania herbacea* (Mill.) McVaugh nodules grown under flooded and non-flooded conditions. *Int J Mol Sci* **20**, 1824.

Kumagai H, Kinoshita E, Ridge RW, Kouchi H, 2006. RNAi knock-down of *ENOD40*s leads to significant suppression of nodule formation in *Lotus japonicus*. *Plant Cell Physiol* **47**, 1102-11.

Kumar N, Lad G, Giuntini E*, et al.*, 2015. Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum*. *Open Biol* **5**, 140133.

Kusakabe S, Higasitani N, Kaneko T*, et al.*, 2020. *Lotus* accessions possess multiple checkpoints triggered by different type III secretion system effectors of the wide-host-range symbiont *Bradyrhizobium elkanii* USDA61. *Microbes Environ* **35**.

Laffont C, Ivanovici A, Gautrat P, Brault M, Djordjevic MA, Frugier F, 2020. The NIN transcription factor coordinates CEP and CLE signaling peptides that regulate nodulation antagonistically. *Nat Commun* **11**, 3167.

Laloum T, Baudin M, Frances L*, et al.*, 2014. Two CCAAT-box-binding transcription factors redundantly regulate early steps of the legume-rhizobia endosymbiosis. *Plant J* **79**, 757-68.

Langfelder P, Horvath S, 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform* **9**, 559.

Langfelder P, Zhang B, Horvath S, 2008. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* **24**, 719-20.

Laporte P, Lepage A, Fournier J*, et al.*, 2014. The CCAAT box-binding transcription factor NF-YA1 controls rhizobial infection. *J Exp Bot* **65**, 481-94.

Laus MC, Logman TJ, Lamers GE, Van Brussel AA, Carlson RW, Kijne JW, 2006. A novel polar surface polysaccharide from *Rhizobium leguminosarum* binds host plant lectin. *Mol Microbiol* **59**, 1704-13.

Lazo GR, Stein PA, Ludwig RA, 1991. A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. *Nat Biotechnol* **9**, 963-7.

Le Gall H, Philippe F, Domon JM, Gillet F, Pelloux J, Rayon C, 2015. Cell wall metabolism in response to abiotic stress. *Plants (Basel)* **4**, 112-66.

Lee KH, Larue TA, 1992. Exogenous ethylene inhibits nodulation of *Pisum sativum* L. cv Sparkle. *Plant Physiol* **100**, 1759-63.

Lengwati DM, Mathews C, Dakora FD, 2020. Rotation benefits from N2-fixing grain legumes to cereals: from increases in seed yield and quality to greater household cash-income by a following maize crop. *Frontiers in Sustainable Food Systems* **4**.

Lewin A, Cervantes E, Chee-Hoong W, Broughton WJ, 1990. *nodSU*, two new nod genes of the broad host range *Rhizobium strain* NGR234 encode host-specific nodulation of the tropical tree *Leucaena leucocephala*. *Mol Plant Microbe Interact* **3**, 317-26.

Li J, Sima W, Ouyang B*, et al.*, 2012. Tomato *SlDREB* gene restricts leaf expansion and internode elongation by downregulating key genes for gibberellin biosynthesis. *J Exp Bot* **63**, 6407-20.

Liang J, Hoffrichter A, Brachmann A, Marin M, 2018. Complete genome of *Rhizobium leguminosarum* Norway, an ineffective *Lotus* micro-symbiont. *Stand Genomic Sci* **13**, 36.

Liang J, Klingl A, Lin YY, Boul E, Thomas-Oates J, Marin M, 2019. A subcompatible rhizobium strain reveals infection duality in *Lotus*. *J Exp Bot* **70**, 1903-13.

Liang Y, Cao Y, Tanaka K*, et al.*, 2013. Nonlegumes respond to rhizobial Nod factors by suppressing the innate immune response. *Science* **341**, 1384-7.

Lie TA. 1981. Environmental physiology of the legume-Rhizobium symbiosis. In *Nitrogen Fixation* **1**, 104-34.

Liebrand TW, Van Den Berg GC, Zhang Z*, et al.*, 2013. Receptor-like kinase SOBIR1/EVR interacts with receptor-like proteins in plant immunity against fungal infection. *Proc Natl Acad Sci U.S.A.* **110**, 10010-5.

Limpens E, Franken C, Smit P, Willemse J, Bisseling T, Geurts R, 2003. LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* **302**, 630-3.

Lin J, Frank M, Reid D, 2020. No Home without Hormones: how plant hormones control legume nodule organogenesis. *Plant Commun* **1**, 100104.

Liu A, Yang Y, Ren X*, et al.*, 2020. Current progress of electrocatalysts for ammonia synthesis through electrochemical nitrogen reduction under ambient conditions. *Chem Sus Chem* **13**, 3766-88.

Liu CJ, 2012. Deciphering the enigma of lignification: precursor transport, oxidation, and the topochemistry of lignin assembly. *Mol Plant* **5**, 304-17.

Liu CW, Breakspear A, Guan D*, et al.*, 2019. NIN acts as a network hub controlling a growth module required for rhizobial infection. *Plant Physiol* **179**, 1704-22.

Liu H, Zhang C, Yang J, Yu N, Wang E, 2018a. Hormone modulation of legume-rhizobial symbiosis. *J Integr Plant Biol* **60**, 632-48.

Liu YH, Jiao YS, Liu LX*, et al.*, 2018b. Nonspecific symbiosis between *Sophora flavescens* and different rhizobia. *Mol Plant Microbe Interact* **31**, 224-32.

Lombardi N, Vitale S, Turra D*, et al.*, 2018. Root exudates of stressed plants stimulate and attract *Trichoderma* soil fungi. *Mol Plant Microbe Interact* **31**, 982-94.

Lopez M, Herrera-Cervera JA, Iribarne C, Tejera NA, Lluch C, 2008. Growth and nitrogen fixation in *Lotus japonicus* and *Medicago truncatula* under NaCl stress: nodule carbon metabolism. *J Plant Physiol* **165**, 641-50.

Lopez-Lara IM, Blok-Tip L, Quinto C*, et al.*, 1996. NodZ of *Bradyrhizobium* extends the nodulation host range of *Rhizobium* by adding a fucosyl residue to nodulation signals. *Mol Microbiol* **21**, 397-408.

Lorite MJ, Estrella MJ, Escaray FJ*, et al.*, 2018. The Rhizobia-*Lotus* Symbioses: deeply specific and widely diverse. *Front Microbiol* **9**, 2055.

Lorteau MA, Ferguson BJ, Guinel FC, 2001. Effects of cytokinin on ethylene production and nodulation in pea (*Pisum sativum*) cv. Sparkle. *Physiol Plant* **112**, 421-8.

Love MI, Huber W, Anders S, 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.

Mabrouk Y, Hemissi I, Salem IB, Mejri S, Saidi M, Belhadj O, 2018. Potential of rhizobia in improving nitrogen fixation and yields of legumes. In *Symbiosis.*

Macho AP, Zipfel C, 2015. Targeting of plant pattern recognition receptor-triggered immunity by bacterial type-III secretion system effectors. *Curr Opin Microbiol* **23**, 14-22.

Madsen EB, Madsen LH, Radutoiu S*, et al.*, 2003. A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* **425**, 637-40.

Madsen LH, Tirichine L, Jurkiewicz A*, et al.*, 2010. The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat Commun* **1**, 10.

Maekawa T, Maekawa-Yoshikawa M, Takeda N, Imaizumi-Anraku H, Murooka Y, Hayashi M, 2009. Gibberellin controls the nodulation signaling pathway in *Lotus japonicus*. *Plant J* **58**, 183-94.

Maekawa-Yoshikawa M, Muller J, Takeda N*, et al.*, 2009. The temperature-sensitive brush mutant of the legume *Lotus japonicus* reveals a link between root development and nodule infection by rhizobia. *Plant Physiol* **149**, 1785-96.

Magori S, Oka-Kira E, Shibata S*, et al.*, 2009. Too Much Love, a root regulator associated with the long-distance control of nodulation in *Lotus japonicus*. *Mol Plant Microbe Interact* **22**, 259-68.

Marczak M, Mazur A, Koper P, Zebracki K, Skorupska A, 2017. Synthesis of rhizobial exopolysaccharides and their importance for symbiosis with legume plants. *Genes (Basel)* **8**, 360.

Marek-Kozaczuk M, Wdowiak-Wrobel S, Kalita M*, et al.*, 2017. Host-dependent symbiotic efficiency of *Rhizobium leguminosarum* bv. *trifolii* strains isolated from nodules of *Trifolium rubens*. *Antonie Leeuwenhoek* **110**, 1729-44.

Marino D, Andrio E, Danchin EG*, et al.*, 2011. A *Medicago truncatula* NADPH oxidase is involved in symbiotic nodule functioning. *New Phytol* **189**, 580-92.

Marsh JF, Rakocevic A, Mitra RM*, et al.*, 2007. *Medicago truncatula NIN* is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/ calmodulin-dependent protein kinase. *Plant Physiol* **144**, 324-35.

Mendrygal KE, Gonzalez JE, 2000. Environmental regulation of exopolysaccharide production in *Sinorhizobium meliloti*. *J Bacteriol* **182**, 599-606.

Mhamdi A, Van Breusegem F, 2018. Reactive oxygen species in plant development. *Development* **145**, 15.

Michiels J, Dombrecht B, Vermeiren N, Xi C, Luyten E, Vanderleyden J, 1998. *Phaseolus vulgaris* is a non-selective host for nodulation. *FEMS Microbiol Ecol* **26**, 193-205.

Minchin FR, Pate JS, 1975. Effects of water, aeration, and salt regime on nitrogen fixation in a nodulated legume – definition of an optimum root environment. *J Exp Bot* **26**, 60-9.

Miwa H, Okazaki S, 2017. How effectors promote beneficial interactions. *Curr Opin Plant Biol* **38**, 148-54.

Miyata K, Kawaguchi M, Nakagawa T, 2013. Two distinct *EIN2* genes cooperatively regulate ethylene signaling in *Lotus japonicus*. *Plant Cell Physiol* **54**, 1469-77.

Montiel J, Nava N, Cardenas L*, et al.*, 2012. A *Phaseolus vulgaris* NADPH oxidase gene is required for root infection by Rhizobia. *Plant Cell Physiol* **53**, 1751-67.

Mulvaney RL, Khan SA, Ellsworth TR, 2009. Synthetic nitrogen fertilizers deplete soil nitrogen: a global dilemma for sustainable cereal production. *J Environ Qual* **38**, 2295-314.

Mun T, Bachmann A, Gupta V, Stougaard J, Andersen SU, 2016. *Lotus* Base: An integrated information portal for the model legume *Lotus japonicus*. *Sci Rep* **6**, 39447.

Murakami E, Cheng J, Gysel K*, et al.*, 2018. Epidermal LysM receptor ensures robust symbiotic signalling in *Lotus japonicus*. *Elife* **7**, e33506.

Murray JD, Karas BJ, Sato S, Tabata S, Amyot L, Szczyglowski K, 2007. A cytokinin perception mutant colonized by *Rhizobium* in the absence of nodule organogenesis. *Science* **315**, 101-4.

Murray MG, Thompson WF, 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* **8**, 4321-5.

Nadeau JA, Sack FD, 2002. Control of stomatal distribution on the *Arabidopsis* leaf surface. *Science* **296**, 1697-700.

Nakabayashi R, Mori T, Saito K, 2014. Alternation of flavonoid accumulation under drought stress in *Arabidopsis thaliana*. *Plant Signal Behav* **9**, e29518.

Naya L, Ladrera R, Ramos J*, et al.*, 2007. The response of carbon metabolism and antioxidant defenses of alfalfa nodules to drought stress and to the subsequent recovery of plants. *Plant Physiol* **144**, 1104-14.

Naylor D, Coleman-Derr D, 2017. Drought stress and root-associated bacterial communities. *Front Plant Sci* **8**, 2223.

Ng JL, Hassan S, Truong TT*, et al.*, 2015. Flavonoids and auxin transport inhibitors rescue symbiotic nodulation in the *Medicago truncatula* cytokinin perception mutant *cre1*. *Plant Cell* **27**, 2210-26.

Nishad R, Ahmed T, Rahman VJ, Kareem A, 2020. Modulation of plant defense system in response to microbial interactions. *Front Microbiol* **11**, 1298.

Nishida H, Ito M, Miura K, Kawaguchi M, Suzaki T, 2020. Autoregulation of nodulation pathway is dispensable for nitrate-induced control of rhizobial infection. *Plant Signal Behav* **15**, 1733814.

Nishida H, Suzaki T, 2018. Two negative regulatory systems of root nodule symbiosis: how are symbiotic benefits and costs balanced? *Plant Cell Physiol* **59**, 1733-8.

Nouwen N, Gargani D, Giraud E, 2019. The modification of the flavonoid naringenin by *Bradyrhizobium* sp. Strain ORS285 changes the nod genes inducer function to a growth stimulator. *Mol Plant Microbe Interact* **32**, 1517-25.

Nukui N, Ezura H, Yuhashi K, Yasuta T, Minamisawa K, 2000. Effects of ethylene precursor and inhibitors for ethylene biosynthesis and perception on nodulation in *Lotus japonicus* and *Macroptilium atropurpureum*. *Plant Cell Physiol* **41**, 893-7.

O'brien José a, Vega A, Bouguyon E*, et al.*, 2016. Nitrate transport, sensing, and responses in plants. *Mol Plant* **9**, 837-56.

Okazaki S, Kaneko T, Sato S, Saeki K, 2013. Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *Proc Natl Acad Sci U.S.A.* **110**, 17131-6.

Okazaki S, Zehner S, Hempel J, Lang K, Gottfert M, 2009. Genetic organization and functional analysis of the type III secretion system of *Bradyrhizobium elkanii*. *FEMS Microbiol Lett* **295**, 88-95.

Okushima Y, Fukaki H, Onoda M, Theologis A, Tasaka M, 2007. ARF7 and ARF19 regulate lateral root formation via direct activation of *LBD*/*ASL* genes in *Arabidopsis*. *Plant Cell* **19**, 118-30.

Oldroyd GE, Downie JA, 2004. Calcium, kinases and nodulation signalling in legumes. *Nat Rev Mol Cell Biol* **5**, 566-76.

Oldroyd GE, Engstrom EM, Long SR, 2001. Ethylene inhibits the Nod factor signal transduction pathway of Medicago truncatula. *Plant Cell* **13**, 1835-49.

Oldroyd GE, Murray JD, Poole PS, Downie JA, 2011. The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev Genet* **45**, 119-44.

Or D, Smets BF, Wraith JM, Dechesne A, Friedman SP, 2007. Physical constraints affecting bacterial habitats and activity in unsaturated porous media – a review. *Adv Water Resour* **30**, 1505-27.

Paco A, Da-Silva JR, Eliziario F, Brigido C, Oliveira S, Alexandre A, 2019. *traG* gene is conserved across *Mesorhizobium* spp. able to nodulate the same host plant and expressed in response to root exudates. *Biomed Res Int* **2019**, 3715271.

Papadopoulos AP, Bar-Tal A, Silber A, Saha UK, Raviv M, 2008. Inorganic and synthetic organic components of soilless culture and potting mixes. In *Soilless Culture.* 505-43.

Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I, 2018. zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. *Gigascience* **7**.

Peleg-Grossman S, Melamed-Book N, Levine A, 2012. ROS production during symbiotic infection suppresses pathogenesis-related gene expression. *Plant Signal Behav* **7**, 409-15.

Peng Z, Liu F, Wang L*, et al.*, 2017. Transcriptome profiles reveal gene regulation of peanut (*Arachis hypogaea* L.) nodulation. *Sci Rep* **7**, 40066.

Penmetsa RV, Cook DR, 1997. A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. *Science* **275**, 527-30.

Penmetsa RV, Uribe P, Anderson J*, et al.*, 2008. The *Medicago truncatula* ortholog of Arabidopsis EIN2, *sickle*, is a negative regulator of symbiotic and pathogenic microbial associations. *Plant J* **55**, 580-95.

Peters NK, Crist-Estes DK, 1989. Nodule formation is stimulated by the ethylene inhibitor aminoethoxyvinylglycine. *Plant Physiol* **91**, 690-3.

Plet J, Wasson A, Ariel F*, et al.*, 2011. MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in *Medicago truncatula*. *Plant J* **65**, 622-33.

Pucciariello C, Boscari A, Tagliani A, Brouquisse R, Perata P, 2019. Exploring legume-rhizobia symbiotic models for waterlogging tolerance. *Front Plant Sci* **10**, 578.

Puppo A, Groten K, Bastian F*, et al.*, 2005. Legume nodule senescence: roles for redox and hormone signalling in the orchestration of the natural aging process. *New Phytol* **165**, 683-701.

Qing G, Ghazfar R, Jackowski ST*, et al.*, 2020. Recent advances and challenges of electrocatalytic N2 reduction to ammonia. *Chem Rev* **120**, 5437-516.

Radutoiu S, Madsen LH, Madsen EB*, et al.*, 2003. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425**, 585-92.

Radutoiu S, Madsen LH, Madsen EB*, et al.*, 2007. LysM domains mediate lipochitin-oligosaccharide recognition and *Nfr* genes extend the symbiotic host range. *EMBO J* **26**, 3923-35.

Ramu SK, Peng HM, Cook DR, 2002. Nod factor induction of reactive oxygen species production is correlated with expression of the early nodulin gene *rip1* in *Medicago truncatula*. *Mol Plant Microbe Interact* **15**, 522-8.

Recourt K, Van Brussel AA, Driessen AJ, Lugtenberg BJ, 1989. Accumulation of a *nod* gene inducer, the flavonoid naringenin, in the cytoplasmic membrane of *Rhizobium leguminosarum* biovar *viciae* is caused by the pH-dependent hydrophobicity of naringenin. *J Bacteriol* **171**, 4370-7.

Reid D, Liu H, Kelly S*, et al.*, 2018. Dynamics of ethylene production in response to compatible Nod factor. *Plant Physiol* **176**, 1764-72.

Reid DE, Heckmann AB, Novak O, Kelly S, Stougaard J, 2016. CYTOKININ OXIDASE/DEHYDROGENASE3 maintains cytokinin homeostasis during root and nodule development in *Lotus japonicus*. *Plant Physiol* **170**, 1060-74.

Rey P, Tarrago L, 2018. Physiological roles of plant methionine sulfoxide reductases in redox homeostasis and signaling. *Antioxidants (Basel)* **7**, 114.

Ribeiro CW, Alloing G, Mandon K, Frendo P, 2015. Redox regulation of differentiation in symbiotic nitrogen fixation. *Biochim Biophys Acta* **1850**, 1469-78.

Rigby H, Clarke BO, Pritchard DL*, et al.*, 2016. A critical review of nitrogen mineralization in biosolids-amended soil, the associated fertilizer value for crop production and potential for emissions to the environment. *Sci Total Environ* **541**, 1310-38.

Rivilla R, Martin M, Lloret J, 2017. What makes rhizobia rhizosphere colonizers? *Environ Microbiol* **19**, 4379-81.

Robinson JT, Thorvaldsdottir H, Winckler W*, et al.*, 2011. Integrative genomics viewer. *Nat Biotechnol* **29**, 24-6.

Roche P, Debelle F, Maillet F*, et al.*, 1991. Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. *Cell* **67**, 1131-43.

Roche P, Maillet F, Plazanet C*, et al.*, 1996. The common *nodABC* genes of *Rhizobium meliloti* are host-range determinants. *Proc Natl Acad Sci U.S.A.* **93**, 15305-10.

Rodriguez-Navarro DN, Dardanelli MS, Ruiz-Sainz JE, 2007. Attachment of bacteria to the roots of higher plants. *FEMS Microbiol Lett* **272**, 127-36.

Rolfe BG, Carlson RW, Ridge RW, Dazzo FB, Mateos PF, Pankhurst CE, 1996. Defective infection and nodulation of clovers by exopolysaccharide mutants of *Rhizobium leguminosarum* bv. *trifolii*. *Func Plant Bio* **23**, 285-303.

Roy S, Liu W, Nandety RS*, et al.*, 2020. Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. *Plant Cell* **32**, 15-41.

Sadowsky MJ, Cregan PB, 1992. The soybean *Rj4* allele restricts nodulation by *Bradyrhizobium japonicum* serogroup 123 strains. *Appl Environ Microbiol* **58**, 720-3.

Saeki K, Kouchi H, 2000. The Lotus Symbiont, *Mesorhizobium loti*: Molecular genetic techniques and application. *J Plant Res* **113**, 457-65.

Safronova VI, Guro PV, Sazanova AL*, et al.*, 2020. Rhizobial microsymbionts of kamchatka oxytropis species possess genes of the type III and VI secretion systems, which can affect the development of symbiosis. *Mol Plant Microbe Interact* **33**, 1232-41.

Saijo Y, Loo EP, 2020. Plant immunity in signal integration between biotic and abiotic stress responses. *New Phytol* **225**, 87-104.

Saito K, Yoshikawa M, Yano K*, et al.*, 2007. NUCLEOPORIN85 is required for calcium spiking, fungal and bacterial symbioses, and seed production in *Lotus japonicus*. *Plant Cell* **19**, 610-24.

Salas A, Tortosa G, Hidalgo-Garcia A*, et al.*, 2019. The hemoglobin Bjgb from *Bradyrhizobium diazoefficiens* controls NO homeostasis in soybean nodules to protect symbiotic nitrogen fixation. *Front Microbiol* **10**, 2915.

Sampedro J, Cosgrove DJ, 2005. The expansin superfamily. *Genome Biol* **6**, 242.

Sanchez-Romera B, Ruiz-Lozano JM, Li G*, et al.*, 2014. Enhancement of root hydraulic conductivity by methyl jasmonate and the role of calcium and abscisic acid in this process. *Plant Cell Environ* **37**, 995-1008.

Sandal N, Jin H, Rodriguez-Navarro DN*, et al.*, 2012. A set of *Lotus japonicus* Gifu x *Lotus burttii* recombinant inbred lines facilitates map-based cloning and QTL mapping. *DNA Res* **19**, 317-23.

Santos R, Herouart D, Sigaud S, Touati D, Puppo A, 2001. Oxidative burst in alfalfa-*Sinorhizobium meliloti* symbiotic interaction. *Mol Plant Microbe Interact* **14**, 86-9.

Sarker U, Oba S, 2020. Nutraceuticals, phytochemicals, and radical quenching ability of selected drought-tolerant advance lines of vegetable amaranth. *BMC Plant Biol* **20**, 564.

Sasaki T, Suzaki T, Soyano T, Kojima M, Sakakibara H, Kawaguchi M, 2014. Shoot-derived cytokinins systemically regulate root nodulation. *Nat Commun* **5**, 4983.

Sato S, Kaneko T, Nakamura Y, Asamizu E, Kato T, Tabata S, 2001. Structural analysis of a *Lotus japonicus* genome. I. Sequence features and mapping of fifty-six TAC clones which cover the 5.4 mb regions of the genome. *DNA Res* **8**, 311-8.

Schaller A, Stintzi A, Rivas S*, et al.*, 2018. From structure to function - a family portrait of plant subtilases. *New Phytol* **218**, 901-15.

Scharwies JD, Dinneny JR, 2019. Water transport, perception, and response in plants. *J Plant Res* **132**, 311-24.

Schauser L, Roussis A, Stiller J, Stougaard J, 1999. A plant regulator controlling development of symbiotic root nodules. *Nature* **402**, 191-5.

Schiessl K, Lilley JLS, Lee T*, et al.*, 2019. *NODULE INCEPTION* recruits the lateral root developmental program for symbiotic nodule organogenesis in *Medicago truncatula*. *Curr Biol* **29**, 3657-68 e5.

Schimel J, Balser TC, Wallenstein M, 2007. Microbial stress-response physiology and its implications for ecosystem function. *Ecology* **88**, 1386-94.

Schmidt JS, Harper JE, Hoffman TK, Bent AF, 1999. Regulation of soybean nodulation independent of ethylene signaling. *Plant Physiol* **119**, 951-60.

Sessions SK, 2013. Genome size. In *Brenner's Encyclopedia of Genetics (Second Edition).* San Diego: Academic Press, 301-5.

Shamseldin A, Velazquez E, 2020. The promiscuity of *Phaseolus vulgaris* L. (common bean) for nodulation with rhizobia: a review. *World J Microbiol Biotechnol* **36**, 63.

Sharma A, Shahzad B, Rehman A, Bhardwaj R, Landi M, Zheng B, 2019. Response of phenylpropanoid pathway and the role of polyphenols in plants under abiotic stress. *Molecules* **24**.

Sharples SC, Nguyen-Phan TC, Fry SC, 2017. Xyloglucan endotransglucosylase/hydrolases (XTHs) are inactivated by binding to glass and cellulosic surfaces, and released in active form by a heat-stable polymer from cauliflower florets. *J Plant Physiol* **218**, 135-43.

Shimada N, Aoki T, Sato S, Nakamura Y, Tabata S, Ayabe S-I, 2003. A cluster of genes encodes the two types of chalcone isomerase involved in the biosynthesis of general flavonoids and legume-specific 5-deoxy (iso) flavonoids in *Lotus japonicus*. *Plant Physiol.* **131**, 941-51.

Shimada N, Sasaki R, Sato S*, et al.*, 2005. A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome. *J. Exp. Bot.* **56**, 2573-85.

Shimada N, Sato S, Akashi T*, et al.*, 2007. Genome-wide analyses of the structural gene families involved in the legume-specific 5-deoxyisoflavonoid biosynthesis of *Lotus japonicus*. *DNA Res* **14**, 25-36.

Singh S, Katzer K, Lambert J, Cerri M, Parniske M, 2014. CYCLOPS, a DNA-binding transcriptional activator, orchestrates symbiotic root nodule development. *Cell Host Microbe* **15**, 139-52.

Singh S, Parniske M, 2012. Activation of calcium- and calmodulin-dependent protein kinase (CCaMK), the central regulator of plant root endosymbiosis. *Curr Opin Plant Biol* **15**, 444-53.

Smit P, Raedts J, Portyanko V*, et al.*, 2005. NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. *Science* **308**, 1789-91.

Smith VH, 2003. Eutrophication of freshwater and coastal marine ecosystems: a global problem. *Environ Sci Pollut Res Int* **10**, 126-39.

Soby S, Bergman K, 1983. Motility and chemotaxis of *Rhizobium meliloti* in soil. *Appl Environ Microbiol* **46**, 995-8.

Solovev YV, Igolkina AA, Kuliaev PO*, et al.*, 2021. Towards understanding Afghanistan pea symbiotic phenotype through the molecular modeling of the interaction between LykX-Sym10 receptor heterodimer and nod factors. *Front Plant Sci* **12**, 642591.

Solovyev V, Kosarev P, Seledsov I, Vorobyev D, 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol* **7 Suppl 1**, S10 1-2.

Somssich M, 2020. Lignification and Oxidative Enzymes: Localization, Localization, Localization! *Plant Physiol* **184**, 554-5.

Sorensen KK, Simonsen JB, Maolanon NN, Stougaard J, Jensen KJ, 2014. Chemically synthesized 58-mer LysM domain binds lipochitin oligosaccharide. *Chembiochem* **15**, 2097-105.

Soyano T, Hayashi M, 2014. Transcriptional networks leading to symbiotic nodule organogenesis. *Curr Opin Plant Biol* **20**, 146-54.

Soyano T, Kouchi H, Hirota A, Hayashi M, 2013. NODULE INCEPTION directly targets *NF-Y* subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. *PLoS Genet* **9**, e1003352.

Soyano T, Shimoda Y, Kawaguchi M, Hayashi M, 2019. A shared gene drives lateral root development and root nodule symbiosis pathways in *Lotus*. *Science* **366**, 1021-3.

Spencer Graves, Piepho H-P, Selzer L, Dorai-Raj WHFS, 2019. *multcompView: Visualizations of Paired Comparisons*. R package version 0.1-8.

Sponsel VM, Hedden P, 2010. Gibberellin biosynthesis and inactivation. In *Plant Hormones.* 63-94.

Sprent JI, 1984. Effects of drought and salinity on heterotrophic nitrogen fixing bacteria and on infection of legumes by rhizobia. In *Advances in Nitrogen Fixation Research.* 295-302.

Sprent JI, Ardley J, James EK, 2017. Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytol* **215**, 40-56.

Sprent JI, James EK, 2007. Legume evolution: where do nodules and mycorrhizas fit in? *Plant Physiol* **144**, 575-81.

Stacey G, 1995. *Bradyrhizobium japonicum* nodulation genetics. *FEMS Microbiol Lett* **127**, 1-9.

Stark JM, Firestone MK, 1995. Mechanisms for soil moisture effects on activity of nitrifying bacteria. *Appl Environ Microbiol* **61**, 218-21.

Stassen MJJ, Hsu SH, Pieterse CMJ, Stringlis IA, 2021. Coumarin communication along the microbiome-root-shoot axis. *Trends Plant Sci* **26**, 169-83.

Stevens CJ, 2019. Nitrogen in the environment. *Science* **363**, 578-80.

Stougaard J, Abildsten D, Marcker KA, 1987. The *Agrobacterium rhizogenes* pRi TL-DNA segment as a gene vector system for transformation of plants. *Mol Gen Genet* **207**, 251-5.

Stracke S, Kistner C, Yoshida S*, et al.*, 2002. A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* **417**, 959-62.

Striker GG, Colmer TD, 2017. Flooding tolerance of forage legumes. *J Exp Bot* **68**, 1851-72.

Striker GG, Izaguirre RF, Manzur ME, Grimoldi AA, 2012. Different strategies of *Lotus japonicus*, *L. corniculatus* and *L. tenuis* to deal with complete submergence at seedling stage. *Plant Biol (Stuttg)* **14**, 50-5.

Sugawara M, Epstein B, Badgley BD*, et al.*, 2013. Comparative genomics of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. *Genome Biol* **14**, R17.

Sugawara M, Takahashi S, Umehara Y*, et al.*, 2018. Variation in bradyrhizobial NopP effector determines symbiotic incompatibility with *Rj2*-soybeans via effector-triggered immunity. *Nat Commun* **9**, 3139.

Sung FJM, 1993. Waterlogging effect on nodule nitrogenase and leaf nitrate reductase activities in soybean. *Field Crops Research* **35**, 183-9.

Suzaki T, Ito M, Kawaguchi M, 2013. Induction of localized auxin response during spontaneous nodule development in *Lotus japonicus*. *Plant Signal Behav* **8**, e23359.

Suzuki A, Akune M, Kogiso M*, et al.*, 2004. Control of nodule number by the phytohormone abscisic acid in the roots of two leguminous species. *Plant Cell Physiol* **45**, 914-22.

Takahara M, Magori S, Soyano T*, et al.*, 2013. Too Much Love, a novel Kelch repeat-containing F-box protein, functions in the long-distance regulation of the legume-*Rhizobium* symbiosis. *Plant Cell Physiol* **54**, 433-47.

Tanabata S, Ohyam T, 2014. Autoregulation of nodulation in soybean plants. In *Advances in Biology and Ecology of Nitrogen Fixation.*

Tang F, Yang S, Liu J, Zhu H, 2016. *Rj4*, a gene controlling nodulation specificity in soybeans, encodes a thaumatin-like protein but not the one previously reported. *Plant Physiol* **170**, 26-32.

Tang H, Bomhoff MD, Briones E, Zhang L, Schnable JC, Lyons E, 2015. SynFind: Compiling syntenic regions across any set of genomes on demand. *Genome Biol Evol* **7**, 3286-98.

Terpolilli JJ, O'hara GW, Tiwari RP, Dilworth MJ, Howieson JG, 2008. The model legume *Medicago truncatula* A17 is poorly matched for N2 fixation with the sequenced microsymbiont *Sinorhizobium meliloti* 1021. *New Phytol* **179**, 62-6.

Tesfaye M, 1998. *Rhizobium* strains that nodulate *Trifolium semipilosum* Fres. are phylogenetically distinct. *Plant and Soil* **207**, 147-54.

Theodoulou FL, Kerr ID, 2015. ABC transporter research: going strong 40 years on. *Biochem Soc Trans* **43**, 1033-40.

Tikhonovich IA, Provorov NA, 2007. Chapter 14 – Beneficial plant-microbe interactions. In *Comprehensive and Molecular Phytopathology.* Amsterdam: Elsevier, 365-420.

Tirichine L, James EK, Sandal N, Stougaard J, 2006. Spontaneous root-nodule formation in the model legume *Lotus japonicus*: a novel class of mutants nodulates in the absence of rhizobia. *Mol Plant Microbe Interact* **19**, 373-82.

Tirichine L, Sandal N, Madsen LH*, et al.*, 2007a. A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science* **315**, 104-7.

Tominaga A, Nagata M, Futsuki K*, et al.*, 2010. Effect of abscisic acid on symbiotic nitrogen fixation activity in the root nodules of *Lotus japonicus*. *Plant Signal Behav* **5**, 440-3.

Tsukui T, Eda S, Kaneko T*, et al.*, 2013. The type III secretion system of *Bradyrhizobium japonicum* USDA122 mediates symbiotic incompatibility with *Rj2* soybean plants. *Appl Environ Microbiol* **79**, 1048-51.

Tsurumaru H, Hashimoto S, Okizaki K, Kanesaki Y, Yoshikawa H, Yamakawa T, 2015. A putative type III secretion system effector encoded by the *MA20_12780* gene in *Bradyrhizobium japonicum* Is-34 causes incompatibility with *Rj4* genotype soybeans. *Appl Environ Microbiol* **81**, 5812-9.

Tucker M, Wen C-K, 2015. Research Tool: Ethylene preparation: Treatment with ethylene and its replacements. In *Ethylene in Plants.* 245-61.

Ullah A, Manghwar H, Shaban M*, et al.*, 2018. Phytohormones enhanced drought tolerance in plants: a coping strategy. *Environ Sci Pollut Res Int* **25**, 33103-18.

Van Zeijl A, Op Den Camp RH, Deinum EE*, et al.*, 2015. Rhizobium lipo-chitooligosaccharide signaling triggers accumulation of cytokinins in *Medicago truncatula* roots. *Mol Plant* **8**, 1213-26.

Varshney RK, Thudi M, Pandey MK*, et al.*, 2018. Accelerating genetic gains in legumes for the development of prosperous smallholder agriculture: integrating genomics, phenotyping, systems modelling and agronomy. *J Exp Bot* **69**, 3293-312.

Venado RE, Liang J, Marín M, 2020. Chapter 4 - Rhizobia infection, a journey to the inside of plant cells. In *Regulation of Nitrogen-Fixing Symbioses in Legumes*. *Adv Bot Res* **94**, 97-118.

Verdonck O, De Vleeschauwer D, De Boodt M, 1982. The influence of the substrate to plant Growth. *Acta Hortic* **126**, 251-8.

Verma AK, Kumar S, Das M, Dwivedi PD, 2013. A comprehensive review of legume allergy. *Clin Rev Allergy Immunol* **45**, 30-46.

Verma V, Ravindran P, Kumar PP, 2016. Plant hormone-mediated regulation of stress responses. *BMC Plant Biol* **16**, 86.

Vernie T, Kim J, Frances L*, et al.*, 2015. The NIN transcription factor coordinates diverse nodulation programs in different tissues of the *Medicago truncatula* root. *Plant Cell* **27**, 3410-24.

Vervliet G, Holsters M, Teuchy H, Van Montagu M, Schell J, 1975. Characterization of different plaque-forming and defective temperate phages in *Agrobacterium*. *J Gen Virol* **26**, 33-48.

Vest G, Caldwell BE, 1972. *Rj4* – A Gene conditioning ineffective nodulation in soybean. *Crop Sci* **12**, 692-3.

Vriezen JA, De Bruijn FJ, Nusslein K, 2007. Responses of rhizobia to desiccation in relation to osmotic stress, oxygen, and temperature. *Appl Environ Microbiol* **73**, 3451-9.

Walker L, Lagunas B, Gifford ML, 2020. Determinants of host range specificity in legume-rhizobia symbiosis. *Front Microbiol* **11**, 585749.

Wallace TC, Murray R, Zelman KM, 2016. The nutritional value and health benefits of chickpeas and hummus. *Nutrients* **8**.

Wang C, Reid JB, Foo E, 2018a. The art of self-control – autoregulation of plant-microbe symbioses. *Front Plant Sci* **9**, 988.

Wang D, Yang S, Tang F, Zhu H, 2012. Symbiosis specificity in the legume: rhizobial mutualism. *Cell Microbiol* **14**, 334-42.

Wang Q, Liu J, Zhu H, 2018b. Genetic and molecular mechanisms underlying symbiotic specificity in legume-rhizobium interactions. *Front Plant Sci* **9**, 313.

Wang Q, Wang L, Chandrasekaran U, Luo X, Zheng C, Shu K, 2021. ABA biosynthesis and signaling cascades under hypoxia stress. *Front Plant Sci* **12**, 661228.

Wani SH, Kumar V, Shriram V, Sah SK, 2016. Phytohormones and their metabolic engineering for abiotic stress tolerance in crop plants. *The Crop Journal* **4**, 162-76.

Webb KJ, Skot L, Nicholson MN, Jorgensen B, Mizen S, 2000. *Mesorhizobium loti* increases root-specific expression of a calcium-binding protein homologue identified by promoter tagging in *Lotus japonicus*. *Mol Plant Microbe Interact* **13**, 606-16.

Wickham H, 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham H, 2019. *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.4.0.

Wickham H, François R, Henry L, Müller K, 2021a. *dplyr: A Grammar of Data Manipulation*. R package version 1.0.7.

Wickham H, Henry L, Vaughan D, 2021b. *vctrs: Vector Helpers*. R package version 0.3.8.

Wickham H, Hester J, 2020. *readr: Read Rectangular Text Data*. R package version 1.4.0.

Wijffelman C, Spaink H, Schlaman H*, et al.*, 1989. Regulation of *Nod* gene expression: the role of Nod D protein. In *Signal Molecules in Plants and Plant-Microbe Interactions*, 137-44.

William S, Feil H, Copeland A, 2012. Bacterial genomic DNA isolation using CTAB. *Sigma* **50**.

Williams RJ, Spencer JP, Rice-Evans C, 2004. Flavonoids: antioxidants or signalling molecules? *Free Radic Biol Med* **36**, 838-49.

Winkel-Shirley B, 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol* **126**, 485-93.

Wong J, Nadzieja M, Madsen LH*, et al.*, 2019. A *Lotus japonicus* cytoplasmic kinase connects Nod factor perception by the NFR5 LysM receptor to nodulation. *Proc Natl Acad Sci U.S.A.* **116**, 14339-48.

Yamaguchi S, Smith MW, Brown RG, Kamiya Y, Sun T, 1998. Phytochrome regulation and differential expression of gibberellin 3beta-hydroxylase genes in germinating Arabidopsis seeds. *Plant Cell* **10**, 2115-26.

Yang H, Li S, Sun H, Wang W, Zhao F, 2018. Effects of substrate material on plant growth and nutrient loss. *Pol J Environ Stud* **27**, 2821-32.

Yang S, Tang F, Gao M, Krishnan HB, Zhu H, 2010. R gene-controlled host specificity in the legume-rhizobia symbiosis. *Proc Natl Acad Sci U.S.A*. **107**, 18735-40.

Yano K, Yoshida S, Muller J*, et al.*, 2008. CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proc Natl Acad Sci U.S.A.* **105**, 20540-5.

You J, Chan Z, 2015. ROS regulation during abiotic stress responses in crop plants. *Front Plant Sci* **6**, 1092.

Zaat SA, Wijffelman CA, Spaink HP, Van Brussel AA, Okker RJ, Lugtenberg BJ, 1987. Induction of the *nodA* promoter of *Rhizobium leguminosarum Sym* plasmid pRL1JI by plant flavanones and flavones. *J Bacteriol* **169**, 198-204.

Zahran HH, 1999. *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol Mol Biol Rev* **63**, 968-89, table of contents.

Zaller JG, Frank T, Drapela T, 2011. Soil sand content can alter effects of different taxa of mycorrhizal fungi on plant biomass production of grassland species. *Eur J Soil Biol* **47**, 175-81.

Zarrabian M, Montiel J, Sandal N*, et al.*, 2021. A promiscuity locus confers *Lotus burttii* nodulation with rhizobia from five different genera. DOI: 10.1101/2021.08.26.457880.

Zdyb A, Demchenko K, Heumann J*, et al.*, 2011. Jasmonate biosynthesis in legume and actinorhizal nodules. *New Phytol* **189**, 568-79.

Zhang B, Wang M, Sun Y*, et al.*, 2021. *Glycine max NNL1* restricts symbiotic compatibility with widely distributed bradyrhizobia via root hair infection. *Nat Plants* **7**, 73-86.

Zhang J, Subramanian S, Stacey G, Yu O, 2009. Flavones and flavonols play distinct critical roles during nodulation of *Medicago truncatula* by *Sinorhizobium meliloti*. *Plant J* **57**, 171-83.

Zhang W, Hu W, Wen CK, 2010. Ethylene preparation and its application to physiological experiments. *Plant Signal Behav* **5**, 453-7.

Zhang W, Wen CK, 2010. Preparation of ethylene gas and comparison of ethylene responses induced by ethylene, ACC, and ethephon. *Plant Physiol Biochem* **48**, 45-53.

Zhou S, Zhang C, Huang Y, Chen H, Yuan S, Zhou X, 2021. Characteristics and research progress of legume nodule senescence. *Plants (Basel)* **10**, 1103.

Zhu JK, 2002. Salt and drought stress signal transduction in plants. *Annu Rev Plant Biol* **53**, 247-73.

Ziegenhain C, Vieth B, Parekh S, Hellmann I, Enard W, 2018. Quantitative single-cell transcriptomics. *Brief Funct Genomics* **17**, 220-32.

Supplementary Tables

Table S1. List of bacterial strains used in this study

Strain	Description	Reference
Mesorhizobium loti MAFF 303099-*Ds*Red	MAFF 3030999 strain expressing *Ds*Red, Gm^R	(Maekawa-Yoshikawa et al., 2009)
Rhizobium leguminosarum Norway-GFP	Spontaneous Sm^R mutant of *Rl* Norway containing pFAJ-GFP plamid, Sm^R, Tc^R	(Liang et al., 2019)
Ensifer fredii HH103	Wild type	(Bellato et al., 1997)
Allorhizobium undicola LMGT	Wild type	(de Lajudie et al., 1998)
Escherichia coli TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*Z ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(araleu)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G, Sm^R	Invitrogen
Agrobacterium rhizogenes AR1193	pRi1193 carrying pBR322 in the TL segment, Rf^R, Cb^R	(Stougaard et al., 1987)
Agrobacterium tumefaciens Agl1	pTiBo542 ΔT, Rf^R, Cb^R	(Lazo et al., 1991)

Table S2. List of plant materials used in this study

Lotus accessions and mutants	Experiment	Seed bag number
L. burttii B-303	Nodulation phenotyping	#92871
	RNA-seq	#92872
	CRISPR-Cas12a gene editing	#92872
L. japonicus Gifu B-129	Nodulation phenotyping	#111213, #111214, #111222
	Trans-complementation assay	#110888, #110889, #110890, #110919, #113441
	RNA-seq	#111213
	Root colonization assay	#111213, #111214
	Naringenin treatment assay	#113441
	ROS staining assay	#93067, #113441
L. japonicus MG-20	Nodulation phenotyping	#110941
L. japonicus MG-70	Nodulation phenotyping	#020015915
L. japonicus MG-86	Nodulation phenotyping	#92888
L. japonicus MG-119	Nodulation phenotyping	#92916
L. japonicus MG-123	Nodulation phenotyping	#020016143
L. japonicus MG-135	Nodulation phenotyping	#020016324
L. japonicus MG-136	Nodulation phenotyping	#020016500
L. corniculatus Leo	Nodulation phenotyping	Georg Andreae GmbH
Lj Gifu-*rlp1-1* (L30138037)	Nodulation phenotyping	#113107, #113111
Lj Gifu-*rlp2-1* (L30072908)	Nodulation phenotyping	#113422, #113445
Lj Gifu-*rlp2-2* (L30121643)	Nodulation phenotyping	#114139, #113098
Lj Gifu-*rlp3-1* (L30068880)	Nodulation phenotyping	#113420, #113460
Lj Gifu-*rlp3-2* (L30103870)	Nodulation phenotyping	#112272, #111993
	Trans-complementation assay	#112272, #111983
Lj Gifu-*rlp3-3* (L30112377)	Nodulation phenotyping	#113427, #113454
Lj Gifu-*snf1-1*	Substrate moisture examination	#90793
Lj Gifu-*snf2-2*	Substrate moisture examination	#91473

Table S3. Insertions of LORE1 lines used in this study

Name	Additional insertions			Primers for genotyping	
	Exonic	Intronic	Intergenic	Primer 1[1,2]	Primer 2[2]
Lj Gifu-*rlp1*					
30138037[3]	1	2	1	RLP1 IF2	RLP1 R1
30107408	4	3	1	RLP1 R3	RLP1 F3
30083226	5	5	5	RLP1 R1	RLP1 IF3
30052724	3	6	0	RLP1 R3	RLP1 F3
30001263	1	2	2	RLP1 F3	RLP1 R3
Lj Gifu-*rlp2*					
30072908[3]	0	1	0	30121643 F	30121643 R
30121643[3]	0	1	0	30121643 F	30121643 F
30137696[3]	0	0	0	30137696 F	30137696 R
30132121	0	2	1	RLP2 R5(G)	RLP2 seq(B)
30079887	4	2	2	30121643 R	30121643 F
30082732	1	4	1	RLP2 F2	RLP2 R3
30137695	1	2	0	30137696 F	30137696 R
30052586	2	0	0	30121643 F	30121643 R
30069964	1	1	0	30121643 F	30121643 R
Lj Gifu-*rlp3*					
30103870[3]	0	4	0	30103870 F2	30103870 R2
30056375[3]	0	1	2	30056375 F	30056375 R
30021050[3]	0	1	1	30021050 F	30021050 R
30068880[3]	1	3	0	30103870 F2	30103870 R2
30112377[3]	1	0	0	30103870 F2	30103870 R2
30078379	2	3	1	30103870 R2	30103870 F2

[1] Primer also used with LORE1-P2 primer for detecting LORE1 insertion

[2] Primer sequences are listed in Table S4

[3] LORE1 lines used in phenotyping experiments

Table S4. Primers and oligos used for genotyping, Sanger sequencing, cloning and qRT-PCR

Name	Nucleotide sequence (5' → 3')	Reference
Genotyping primers		
TM1203 F	TTGAATAAGGCTCATAGATCC	(Sato et al., 2001)
TM1203 R	CTTCAGTTTGGGTTTCAAGC	(Sato et al., 2001)
RLP1 R1	AGAAAGCTGAGGTAGATGAA	This study
RLP1 F3	ACCATTACCCCAGTCTTT	This study
RLP1 R3	GCTCTGTTGTTTGTCCTT	This study
RLP1 IF2	CTCGACCTTAGTTGCAGTC	This study
RLP1 IF3	TCTCCATTGTACTCTAAAA	This study
RLP2 R5(G)	TTATATCCTCTTGACTCCATTT	This study
RLP2 seq(B)	GCACGACTTCAGAATCCGCAAG	This study
RLP2 F2	CCTTCCACCATTCTTATTAC	This study
RLP2 R3	AGAGCCAAAAAAGTCATTAT	This study
30121643 F	GCCCATTCAGTTCACACCCCGAAA	Lotus Base
30121643 R	CAGATTGCTGCGAGTGGGATGGTG	Lotus Base
30137696 F	TGACCAGCTGAGAGGTCAGCTTCCAA	Lotus Base
30137696 R	TGGCAAAGAATAACACCAATGCGGA	Lotus Base
30103870 F2	TCCTGGAGGGAACATCCCCGGTTA	Lotus Base
30103870 R2	TGTTCCTTTGCCATGTTTCTCCTTCCA	Lotus Base
30056375 F	TCAGCTTCCAAAGTCCAACTGGAGCA	Lotus Base
30056375 R	TGATCCTGTGAGTTGGTTGTCACCGA	Lotus Base
30021050 F	TTTCAGTGGCCCTCTGCCAGCATC	Lotus Base
30021050 R	AGCATCCCAAACACAGCCCCACAC	Lotus Base
LORE1-P2	CCATGGCGGTTCCGTGAATCTTAGG	Lotus Base
RLP1-FF_CRISPR	TTTGGCAAGTGAAGTCGT	This study
RLP1-FR_CRISPR	AAAGACTGGGGTAATGGT	This study
RLP2-FF_CRISPR	ACAACCTAAGAGGCCAGA	This study
RLP2-FR_CRISPR	GAACAAGTGATGACGCCA	This study
RLP3-EF_CRISPR	TAGCACCATCTTCCAGCTGAGG	This study
RLP3-ER_CRISPR	GTGACAATAGAGAGATCCAAGTACC	This study
RLP4-FF_CRISPR	CAGGGCCAATTAGTGAA	This study
RLP4-FR_CRISPR	TGAAGCGGTGCTAAGA	This study
Genome resequencing primers		
seq R1	GGCCCAATGGGATCTGGTT	This study
seq F2	ACCAGAATATCAAGGTCGGA	This study
seq R2	GAGCTTTATCGCTCCCAAGC	This study
seq F3	GCTTGGGAGCGATAAAGCTC	This study
seq R1f1	CGACAATCATAGCCGTCAGT	This study
seq R1r1	TCACGATTTAGTGGTTCAATCG	This study

seq F2f1	CTTCCGAGCTATGATTAGCCC	This study
seq F2r1	GGAAGAGTATCATAGATGAGAACTGG	This study
seq R2r1	AAGCCAAGGGGTAGATGAAC	This study
seq F3r1	GAGATTAGCGTGCCAGTTGT	This study
seq R2r2	ACTGCTTGATCATAGTTCCAAC	This study
Cloning primers		
LbRLP4 OF1	ATGAAGACATTACGGGTCTCACACCATGGGGTGGTTTCCTTTACCATA	This study
LbRLP4 IR1	ATGAAGACATGTCCAGACCAATCACATGG	This study
LbRLP4 IF2	ATGAAGACATGGACCTTAGTTGCGATCATCTT	This study
LbRLP4 IR2	ATGAAGACATAGAGCCAGAAAAATAATTAAAACCAA	This study
LbRLP4 IF3	ATGAAGACATCTCTCCATTGTATTCTGGAATTG	This study
LbRLP4 IR3	ATGAAGACATGGATGACATGTCTATGCCATCAA	This study
LbRLP4 IF4	ATGAAGACATATCCATCAAGGTAAACTCTTTCAT	This study
LbRLP4 OR4	ATGAAGACATCAGAGGTCTCACCTTTCTGTGGTTTGTACGGGCTCT	This study
pLbRLP4mi F	ATGAAGACATTACGGGTCTCAGCGGATCATATTAAAATAACCACTTTA	This study
pLbRLP4 R	ATTTTTTCCCATGAAGACATCAGAGGTCTCACAGAGTGAAAAATAGAAGAAGCTATATGGAGAAT	This study
gRNAs and oligos for CRISPR-Cas12a editing		
gRLP1.1	ATGAGAAGCTCTTACACGCAT	This study
gRLP1.2	TGGAACTCCACAAGGCCACTC	This study
gRLP2.1	AGGGAGCTTCATCTAGATCTT	This study
gRNA2.2	TAGCCAACTGGGAAATGGATC	This study
gRLP3	ATCCAACAACATGGAAAAAAC	This study
gRLP4.1	TCCAAGACTCTACCTTAGAAG	This study
gRLP4.2	TCTCTAACAACAAACTGCAAG	This study
gRLP4.3	ATGACAAGTCTAATTCAGTAA	This study
gRLP4.4	ATGAGAAGCTCTTACACTCAT	This study
Oligo 1-1A	ATGAAGACTTTACGGGTCTCAAGATAGGGAGCTTCATCTAGATCTTAATTTCTACTAAGTGTTGAGACCTCTGAAGTCTTCAT	This study
Oligo 1-1B	ATGAAGACTTTACGGGTCTCAGTGTAGATATGAGAAGCTCTTACACGCATAATTTCTACTAAGTGTAGATATCCAACAACATGGAAAAAACAATTTGAGACCTCTGAAGTCTTCAT	This study
Oligo 1-2A	ATGAAGACTTTACGGGTCTCAAGATTAGCCAACTGGGAAATGGATCAATTTCTACTAAGTGTTGAGACCTCTGAAGTCTTCAT	This study

Oligo 2-1A	ATGAAGACTTTACGGGTCTCAAGATATCCAACAACATG GAAAAAACAATTTCTACTAAGTGTAGATTGGAACTCCA CAAGGCCACTCAATTTGAGACCTCTGAAGTCTTCAT	This study
Oligo 2-2A	ATGAAGACTTTACGGGTCTCAAGATATCCAACAACATG GAAAAAACAATTTCTACTAAGTGTAGATATGAGAAGCT CTTACACGCATAATTTGAGACCTCTGAAGTCTTCAT	This study
Oligo 3-1A	ATGAAGACTTTACGGGTCTCAAGATTCCAAGACTCTAC CTTAGAAGAATTTCTACTAAGTGTAGATTCTCTAACAA CAAACTGCAAGAATTTGAGACCTCTGAAGTCTTCAT	This study
Oligo 3-2A	ATGAAGACTTTACGGGTCTCAAGATATGACAAGTCTAA TTCAGTAAAATTTCTACTAAGTGTAGATATGAGAAGCT CTTACACTCATAATTTGAGACCTCTGAAGTCTTCAT	This study
qPCR primers		
ATP F	CAATGTCGCCAAGGCCCATGGTG	(Kawaharada et al., 2015)
ATP R	AACACCACTCTCGATCATTTCTCTG	
NIN qF	AACTCACTGGAAACAGGTGCTTTC	(Kumagai et al., 2006)
NIN qR	CTATTGCGGAATGTATTAGCTAGA	
RLP1 qF	GCGGCTGGTTTCAATGTTCCT	This study
RLP1 qR	AATCACGTGGCCTAACATGG	This study
RLP2 qF	TTAGTTGCGGCCATCTTCA	This study
RLP2 qR	CAGATGTGTGAGATAGAAGAGGTT	This study
RLP3 qF	TGTTGACGGCAGTTTTCCA	This study
RLP3 qR	GTTCCAcGACTGTGAGAGCTT	This study
RLP4 qF	CGATTCCACAATGCCTATGCAA	This study
RLP4 qR	TGTGGAATCGTGCCAGTCAA	This study
16Srrna_RlNorway_fw	GTCCATTACTGACGCTGAGG	Sara Masachis
16Srrna_RlNorway_rv	CCCGAAGGGAACCTTGCA	Gelo

[1]BpiI recognition sites are depicted in blue.

[2]BsaI recognition sites are depicted in orange.

[3]gRNA sequences are underlined.

Table S5. List of constructs used in this study

Name	Description
Trans-complementation assay	
LI *LbRLP2*$_{pro}$[1]	Plasmid containing the promoter sequence of *LbRLP2*, Gm^R
LI *LbRLP2*[1]	Plasmid containing the genomic sequence of *LbRLP2*, Gm^R
LI *LbRLP4*$_{pro}$	Plasmid containing the promoter sequence of *LbRLP4*, Gm^R
LI *LbRLP4*	Plasmid containing the genomic sequence of *LbRLP4*, Gm^R
LII 1-2 *35S*$_{pro}$*:NLS:2XmCherry*	LII construct expressing transformation marker
LIII 1-2 *Ub10*$_{pro}$*:NLS:2XmCherry* 3-4 dy[2]	Empty vector with transformation marker; Kn^R
LIII 1-2 *Ub10*$_{pro}$*:NLS:2XmCherry* 3-4 *LbRLP2*$_{pro}$*:LbRLP2:HA*	*LbRLP2* under the control of *LbRLP2* native promoter and transformation marker; Kn^R
LIII 1-2 *35S*$_{pro}$*:NLS:2XmCherry* 3-4 dy	Empty vector with transformation marker; Kn^R
LIII 1-2 *35S*$_{pro}$:NLS:2XmCherry 3-4 *AtUb10*$_{pro}$*:LbRLP2:HA*	*LbRLP2* under the control of *AtUb10* promoter and transformation marker; Kn^R
LIII 1-2 *35S*$_{pro}$*:NLS:2XmCherry* 4-5 *LjUbi*$_{pro}$*-LbRLP4:6XHis*	*LbRLP4* under the control of *LjUbi* promoter and transformation marker; Kn^R
LIII 1-2 *35S*$_{pro}$*:NLS:2XmCherry* 3-4 *AtUb10*$_{pro}$*:LbRLP2:HA* 45 *LjUbi*$_{pro}$*:LbRLP4:6XHis*	*LbRLP2* and *LbRLP4* under the control of *AtUb10* and *LjUbi* promoters, and transformation marker; Kn^R
LIII 1-2 *35S*$_{pro}$*:NLS:2XmCherry* 3-4 *LbRLP2*$_{pro}$*:LbRLP2:HA*	*LbRLP2* under the control of *LbRLP2* native promoter, and transformation marker; Kn^R
LIII 1-2 *35S*$_{pro}$*:NLS:2XmCherry* 4-5 *LbRLP4*$_{pro}$*:LbRLP4:6XHis*	*LbRLP4* under the control of *LbRLP4* native promoter and transformation marker; Kn^R
LIII 12 *35S*$_{pro}$*:NLS:2XmCherry* 3-4 *LbRLP2*$_{pro}$*-LbRLP2:HA* 45 *LbRLP4*$_{pro}$*:LbRLP4:6XHis*	*LbRLP2* and *LbRLP4* under the control of *LbRLP2* and *LbRLP4* native promoters, and transformation marker; Kn^R
CRISPR-Cas12a gene editing assay	
LI 1-2 *LjU6*$_{pro}$[3]	Plasmid containing the *LjU6* promoter, Gm^R
LI 1-2 *LjU6*-T[3]	Plasmid containing the *LjU6* terminator, Gm^R
LII 1-2 *Nos*$_{pro}$*:Hyg*^{R} [3]	Hygromycine resistant gene under the control of *Nos* promoter for transformed plant selection
LII 2-3 *LjUb10*$_{Pro}$*:ttCas12a*[3]	*Cas12a* under the control of *LjUb10* promoter for gene editing
LIII 1-2 *Nos*$_{pro}$*:Hyg*^{R} 2-3 *LjUb10*$_{Pro}$*:ttCas12a* 4-5 *LjU6*$_{pro}$*:gRLP123* Y1	Three gRNAs targeting *RLP1, RLP2* and *RLP3*; Kn^R

LIII 1-2 *Nos$_{pro}$:Hyg^{R}* 2-3 *LjUb10$_{Pro}$:ttCas12a* 4-5 *LjU6$_{pro}$:gRLP123* Y2	Three gRNAs targeting *RLP1, RLP2* and *RLP3*; Kn^{R}
LIII 1-2 *Nos$_{pro}$:Hyg^{R}* 2-3 *LjUb10$_{Pro}$:ttCas12a* 4-5 *LjU6$_{pro}$:gRLP13* Y1	Two gRNAs targeting *RLP1* and *RLP3*; Kn^{R}
LIII 1-2 *Nos$_{pro}$:Hyg^{R}* 2-3 *LjUb10$_{Pro}$:ttCas12a* 4-5 *LjU6$_{pro}$:gRLP13* Y2	Two gRNAs targeting *RLP1* and *RLP3*; Kn^{R}
LIII 1-2 *Nos$_{pro}$:Hyg^{R}* 2-3 *LjUb10$_{Pro}$:ttCas12a* 4-5 *LjU6$_{pro}$:gRLP4* Y1	Two gRNAs targeting *RLP4*; Kn^{R}
LIII 1-2 *Nos$_{pro}$:Hyg^{R}* 2-3 *LjUb10$_{Pro}$:ttCas12a* 4-5 *LjU6$_{pro}$:gRLP4* Y2	Two gRNAs targeting *RLP4*; Kn^{R}

NLS, nuclear localization signal; gRLP, guide RNA for selected RLP encoding gene; Gm^{R}, gentamycin resistance; Kn^{R}, kanamycin resistance; Hyg^{R}, hygromycin resistance

[1, 2] Constructs made by M. Zhang (2016) and L. Eccleston (2017), respectively

[3] Constructs provided by M. Bircheneder (personal communication)

Table S6. Composition of reagents used in CRISPR-Cas12a genome editing assay

Reagent	Preparation	Sterilization	Storage
Potassium phosphate buffer, 0.3 M, pH 6.8	K_2HPO_4/KH_2PO_4 in distilled water	autoclaved	room temperature
Gamborg's B5 medium	in distilled water	autoclaved	room temperature
Gamborg's vitamin 1000X (B5 vitamin)	in sterile water	by 0.22 µm filter	-20°C
1M MES buffer, pH5.2	in sterile water	by 0.22 µm filter	-20°C
BAP[1] (1 µg $µl^{-1}$)	in 1M NaOH	by 0.22 µm filter	-20°C
NAA[2] (1 µg $µl^{-1}$)	in 96% ethanol	by 0.22 µm filter	-20°C
4 M $(NH_4)_2SO_4$	in sterile water	by 0.22 µm filter	-20°C
Hygromycin B (50 µg $µl^{-1}$)	in sterile water	by 0.22 µm filter	-20°C
Cefotaxime (300 µg $µl^{-1}$)	in sterile water	by 0.22 µm filter	-20°C
Acetosyringone (20 µg $µl^{-1}$)	in DMSO	by 0.22 µm filter	-20°C

[1]BAP, 6-benzylaminopurine

[2]NAA, 1-naphthaleneacetic acid

Table S7. Composition of media used in CRISPR-Cas12a genome editing assay

Medium	Composition
YMB liquid medium (100ml) [1]	Mannitol 0.2 g, yeast extract 0.04 g, $MgSO_4.7H_2O$ 0.02 g, NaCl 0.01 g, 0.3 M potassium phosphate buffer, pH7.0 10 µl
B5 medium (100 ml, 1X) [2]	Gamborg B5 basal salts 0.33 g
Co-cultivation medium [2]	0.1X B5 medium, 0.1X B5 vitamin, BAP 0.5 µg ml^{-1}, NAA 0.05 µg ml^{-1}, Acetosyringone 20 µg ml^{-1}, 5 mM MES
Callus induction medium [2]	1X B5 medium, 0.1X B5 vitamin, 2% sucrose, BAP 0.5 µg/ml, NAA 0.05 µg ml^{-1}, 10 mM $(NH_4)_2SO_4$, cefotaxime 300 µg ml^{-1}, hygromycin B 40 µg ml^{-1}, 0.3% Gelrite
Shoot Induction medium [2]	1X B5 medium, 1x vitamin, 2% sucrose, BAP 0.5 µg ml^{-1}, NAA 0.05 µg ml^{-1}, 10 mM NH_4, 0.3% Gelrite
Shoot Elongation medium [2]	1x B5 medium, 1x vitamin, 2% sucrose, BAP 0.2 µg ml^{-1}, 0.3% Gelrite
Root induction medium [2]	0.5X B5 medium, 0.5X vitamin, 1% sucrose, NAA 0.5 µg ml^{-1}, 0.4% Gelrite
Root elongation medium [2]	0.5X B5 medium, 0.5X vitamin, 1% sucrose, 0.5% Gelrite

[1] Reference: (Handberg et al, 1994)

[2] Medium composition from M. Bircheneder, personal communication

Table S8. List of analyzed genes in phenylpropanoid biosynthesis pathway

Enzyme/Function	Gene ID
General phenylpropanoid biosynthesis	
Phenylalanine ammonia lyase [1]	LotjaGi5g1v0272900 LotjaGi1g1v0690200 LotjaGi1g1v0690300 LotjaGi5g1v0102000 LotjaGi5g1v0080000 LotjaGi6g1v0093700 LotjaGi3g1v0071500 LotjaGi1g1v0004200 LotjaGi1g1v0690400_LC
Cinnamic acid 4-hydroxylase [1]	LotjaGi1g1v0106100 LotjaGi5g1v0341100 LotjaGi6g1v0044200
4-coumarate:CoA ligase [1]	LotjaGi3g1v0342800 LotjaGi2g1v0450000 LotjaGi4g1v0299500
Flavonoid biosynthesis	
Chalcone synthase [1]	LotjaGi1g1v0635000* LotjaGi2g1v0449300 LotjaGi3g1v0078700 LotjaGi1g1v0131500 LotjaGi1g1v0268400 LotjaGi1g1v0470600 LotjaGi1g1v0470800 LotjaGi1g1v0599400 LotjaGi2g1v0122800 LotjaGi2g1v0123200_LC LotjaGi2g1v0446700 LotjaGi2g1v0446800 LotjaGi2g1v0447000 LotjaGi2g1v0449400 LotjaGi2g1v0450100 LotjaGi2g1v0451900 LotjaGi3g1v0532600 LotjaGi4g1v0368400 LotjaGi5g1v0136900_LC LotjaGi6g1v0175400
Chalcone isomerase [2]	LotjaGi5g1v0359350 LotjaGi5g1v0359200
Flavanone 3-hydroxylase [1]	LotjaGi4g1v0127900* LotjaGi3g1v0300200*
Flavonol synthase [1]	LotjaGi1g1v0154100* LotjaGi1g1v0327900* LotjaGi1g1v0256000
Polyketide reductase [1] (syn. chalcone reductase)	LotjaGi1g1v0006200 LotjaGi2g1v0187400* LotjaGi6g1v0022500 LotjaGi1g1v0006500
2-hydroxyisoflavanone synthase [1]	LotjaGi4g1v0063000 LotjaGi4g1v0063900

	LotjaGi4g1v0063800
2-hydroxyisoflavanone 4'-O-methyltransferase [3]	LotjaGi4g1v0062900
2-hydroxyisoflavanone dehydratase [1]	LotjaGi5g1v0316500
Isoflavone reductase [1]	LotjaGi2g1v0388100
Vestitone reductase [1]	LotjaGi1g1v0461900 LotjaGi1g1v0462000
Pterocarpan reductase [1]	LotjaGi3g1v0509100 LotjaGi3g1v0509300_LC LotjaGi1g1v0420100
Dihydroflavonol 4-reductase [4]	LotjaGi5g1v0013700 LotjaGi5g1v0013400
Coumarin biosynthesis [5]	
Hydroxycinnamol-Co A shikimate	LotjaGi3g1v0003100 LotjaGi3g1v0347200
Caffeoyl CoA 3-O-methyltransferase	LotjaGi4g1v0076500 LotjaGi4g1v0103100 LotjaGi4g1v0103800
Feruloyl-CoA 6′-Hydroxylase 1	LotjaGi3g1v0076200
Coumarin synthase	LotjaGi6g1v0325100

* Differentially regulated genes
[1] Reference: (Garcia-Calderon et al., 2015, Shimada et al., 2007)
[2] Reference: (Shimada et al., 2003)
[3] Reference: (Akashi et al., 2003)
[4] Reference: (Shimada et al., 2005)
[5] Reference: (Stassen et al., 2021)

List of Tables

List of Figures

Declaration of contribution of research

Yen-Yu Lin, the author of this thesis, contributed to the experiments displayed as follows:

- Figure 1-33: The preparation of the corresponding figures and statistical analysis were done by Yen-Yu Lin.
- Figure 5-8, 11-16, 21-26, 28, 29, 31: Experiments were designed, performed, and analyzed by Yen-Yu Lin.
- Figure 4, 9,10, 17-20, 27, 30, 32: Experiments were designed by Yen-Yu Lin, performed by Yen-Yu Lin with help from collaborator, bachelor or master students, and analyzed by Yen-Yu Lin. The contribution of collaborator and students is listed below.

Raw data acquired from collaborator:

In Figure 10, Lucas Wange constructed the cDNA library for transcriptomic sequencing, mapped the raw reads against the reference genome and examined the quality of transcriptome sequencing.

Raw data acquired from supervising bachelor student thesis of Dorothea Grillmayer:

- In Figure 4, Dorothea Grillmayer inspected the nodulation events in *L. corniculatus*, *L. japonicus* MG-20 and MG-70.
- In Figure 27, Dorothea Grillmayer identifed the SNPs in the 12 candidate genes and re-annotated the QTL together with Yen-Yu Lin.

Raw data acquired from supervising master student theses:

- In Figure 9, 17-20, experiments were designed by Yen-Yu Lin and conducted by Fan-Yu Yu.
- In Figure 30, Entela Malkaj phenotyped L30121643 and Furkan Tunç phenotyped L30068880 and 20112377.
- In Figure 32, Yen-Yu Lin designed the experiments and cloned the constructs. Entela Malkaj performed the *L. burttii* transformation. Entela Malkaj and Yen-Yu Lin performed the plant regeneration. Entela Malkaj and Fan-Yu Yu performed genotyping.

Acknowledgments

Throughout my doctoral study, I have received a great deal of support and assistance.

First of all, I would like to thank my supervisor Dr. Macarena Marín with the greatest appreciation for her invaluable guidance and support through the study. You always gave feedbacks on experimental works and scientific writings. I am thankful that you gave me the opportunity, digging into the topic of moisture-dependent nodulation.

I would like to express my deep appreciation to Prof. Dr. Martin Parniske, Prof. Dr. Jochen Wolf, and Dr. Silke Robatzek for the insightful discussions and valuable advice. My deep gratitude goes to the defense committee members: Prof. Dr. Wolfgang Frank, Prof. Dr. Silke Werth, and Prof. Dr. Christof Osman, for the time and effort to evaluate this work. I would like to extend my great gratitude to Prof. Dr. Wolfgang Enard and Lucas Wange for the collaboration in transcriptome sequencing and associate professor Stig U. Andersen for providing the genome sequences of *L. burttii* and *L. japonicus* Gifu.

I am very grateful to all people who made “Genetics” a wonderful environment. Many thanks to Martin Bircheneder for kindly providing the ‘Stable transformation of *Lotus’* protocol, sharing his experience, and giving practical suggestions. Many thanks to Karl-Heinz Braun, Adriana Hörmann, and Gisela Brinkmann for your technical and moral supports. Many thanks to Isabel Seidler, Nimisha Panchakshari, and Florian Dunker for your help, our gather-together events, and informative conversations. It is my pleasure to have you around. Many thanks to Fang-Yu Hwu and Xiaoyun Gong, Chloé Cathebras, Constance Tisserant, Philipp Bellon, Bernhard Lederer, and An-Po Cheng for the lunch company, bouldering together, ‘Werewolf night’, birthday celebration, and parties. My life gets its color with you. Special thanks to Florian Dunker for your time and effort to give constructive and critical comments on my thesis and discuss with me, Katarzyna Rybak for the supports and discussions on writing strategy, and Kate Parys for the inspirational conversations. I learned a lot.

I am grateful to all members of the Marín group. I would like to extend my gratitude to Juan Liang for giving constructive comments on my thesis, Duncan Crosbie and Dr. Sara Masachis Gelo for correcting part of the English mistakes, and Rafael Venado for the discussions on R programming. Many thanks to my students Chia-Min Liao, Entela Malkaj, Furkan Tunç, Dorothea Grillmayer, and Fan-Yu Yu for the time we worked and enjoyed together. I like the moment we had in prep lab.

Thanks also to all my friends in Munich for their support and company. Special thanks to my first friend in UniChor, Elisabeth Hartl, for your celebration ideas and the time we spent together. Many thanks to Pey-Shin Wang and Lingyun Guo for exploring life in Munich and all the helps. Moreover, I would like to express my sincere appreciation to Ina Ge for your wise advice.

I would like to express my sincerest appreciation to my dearest family and friends, who always have faith in me. Special thanks to Bo-Han Hou for the practical advice on the transcriptome analysis and answered my question in detail. A big thanks to my parents, who have given me the freedom to pursue my goals. A big thanks also to my brother, who handled my documents in Taiwan. My doctoral study would not have been possible without the encouragement and support from you. Even though the Covid-19 pandemic has separated us apart for years, our hearts will always be with each other. Special thanks to my little roommates, Coffee and Rolly, for the time we spent together. You were my greatest support. < May/June 2022, R. I. P. >

Finally, I wish to express my deepest appreciation to my beloved uncle, Prof. Emeritus Dr. Kuang Lin, for being thoughtful and taking care of every family member. I wish I were there by your side. You have set a great example for me. Your dedication in the field of shipping and transportation management will never be forgotten. < July 2021, R. I. P. >

Ina said that everyone has their luck. Thanks to you all, I have become stronger and overcome the encountered difficulties. You are my greatest luck. Thank you very much. Vielen Dank. Muchas gracias. Merci beaucoup. 萬分感謝。