

Université Fédérale



Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Benoit DAUBECH

Le 21 octobre 2019

**Évolution expérimentale d'un symbiote de légumineuse: Étude
des facteurs génétiques et des forces de sélection qui favorisent
ou non l'évolution du mutualisme**

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et
Bioingenieries**

Spécialité : **Interactions plantes-microorganismes**

Unité de recherche :

LIPM - Laboratoire des Interactions Plantes-Microorganismes

Thèse dirigée par

Catherine MASSON-BOIVIN et Marta MARCHETTI

Jury

M. Matthieu Arlat, Professeur, Université Paul Sabatier, Toulouse

M. Benoit Alunni, Maître de conférences, Université Paris Sud, Paris

M. Gilles Béna, Directeur de recherche, IPME, Montpellier

M. Philippe Normand, Directeur de recherche, CNRS, Villeurbanne

Président

Rapporteur

Rapporteur

Examinateur

Mme Marta Marchetti, Co-directrice de thèse

Mme Catherine Masson-Boivin, Directrice de thèse

Remerciements :

Je tiens avant tout à remercier **Claude Bruand** de m'avoir accueilli au LIPM et de m'avoir laissé l'opportunité de m'impliquer dans la vie du laboratoire.

Je tiens aussi à remercier l'ensemble du jury, **Gilles Bena**, **Benoît Alunni**, **Philippe Normand** et **Matthieu Arlat** d'avoir accepté d'évaluer ce travail de thèse.

Je remercie également **Benjamin Gourion**, **Philippe Remigi** et **Delphine Capela** qui ont participé à mes comités de thèses et dont les conseils m'ont beaucoup apporté au cours de ces trois années.

Évidemment, j'adresse de sincères remerciements à **Catherine Masson** et **Marta Marchetti** pour m'avoir encadré, soutenu, encouragé, écouté, secoué, consolé, conseillé et bien plus encore pendant trois ans. Je ne suis pas certain d'avoir été à la hauteur de la tâche tout du long, mais sachez que je vous suis mille fois reconnaissant de m'avoir permis de vivre cette aventure dont je ne garde que le meilleur.

Je tiens également à remercier l'ensemble de l'équipe CMDC (aka JBCM à mon arrivée en stage puis CM) dont j'ai eu le bonheur de partager le quotidien pendant presque quatre années. Je remercie tout particulièrement **Delphine** qui m'a accepté en stage de Master 2 et qui m'a transféré sa passion pour l'expérience d'évolution et sans qui je ne serais pas là. Je remercie aussi **Lukas**, **Mathilda** et **Caroline** de m'avoir soutenu et accompagné pendant cette folle épopée. Un grand merci à **Philippe**, **Anne-Marie** et **Jacques** pour tous les conseils (scientifiques ou non) qu'ils ont pu me donner.

Je tiens aussi à remercier l'ensemble du LIPM où j'ai passé la plus grande partie de mon temps durant ces trois ans de thèse, et où j'ai pu vivre une expérience extraordinaire entouré de gens formidables. Je pense notamment à toute l'équipe SIX (et plus particulièrement à **Aude** et **Manuel**) qui a été un véritable pilier et à **Laurent Sauviac** pour ses précieux conseils de clonage et tout le reste. Je pense aussi à **Fabienne** et **Chrystel** qui ont été des amies fantastiques qui ont toujours su trouver les bons mots et sans qui le temps aurait pu, parfois, sembler long. Il y a aussi **Elie**, **Fernanda**, **Fernando**, **Ariane**, **Claire** et tellement d'autres. Merci aussi à la laverie pour leur disponibilité et leur aide précieuse.

Je ne pourrais oublier l'AJS, dont j'ai eu le plaisir d'être le président pendant un an et grâce à qui j'ai pu faire de belles rencontres avec qui j'ai partagé des moments forts tant au niveau scientifique que personnel.

J'exprime également ma gratitude auprès de tous nos collaborateurs, qui ont permis à cette thèse d'être ce qu'elle est, principalement **Véréna Poinot**, pour tout le temps qu'elle nous a accordé à mes cultures et moi, mais aussi l'équipe de **Sébastien Fort** qui ont fait tout leur possible pour décoder les secrets enzymatiques de pRalta_0471 (aka NoeM) et bien évidemment **Lionel Moulin** et **Agnieszka Klonowska** pour leur contribution sur l'étude de *noeM* ainsi que la plateforme d'imagerie cellulaire pour leur patience et leur gentillesse.

Un grand merci à mes amis, qui ont toujours été présents pour célébrer les bonnes et moins bonnes nouvelles autour d'un verre tout en écoutant la douce et mélodieuse voix de Madonna. Je pense notamment à **Sarah S., Caroline, Chantal, Chloé, Florian, Elodie, Isabelle, Samuel, Iris, Chloé Clémence, Aurore, Charlotte, Guillaume, Marie-Aurore**, et bien évidemment les "MABS" (**Marine, Mikaël, Coco, Alex, Paulo, Paul, Mathieu, Joris**) avec qui j'ai passé deux ans de Master inoubliables et bien plus encore.

Je tiens à dire merci aussi à **Yannick Lerrant** et **Valérie Burtet-Sarramégna** qui ont toujours été là pour me guider depuis mes premiers jours à l'Université de la Nouvelle-Calédonie, et à qui je dois beaucoup.

Pour finir, je tiens à remercier ma famille, mes parents, **Éric** et **Sylvie**, qui ont su me soutenir malgré les milliers de kilomètres qui nous séparent. Vous m'avez toujours encouragé à me surpasser, merci infiniment d'avoir toujours répondu présent, par téléphone ou en personne, quand il le fallait. Je pense aussi à ma sœur **Amélie**, mon frère **Clément** et ma grand-mère **Monique** qui sont toujours là pour me rappeler que tout va bien se passer.

Résumé :

La symbiose qui s'établit entre les légumineuses et les bactéries appelées rhizobia est un processus complexe qui aboutit à la formation d'un nouvel organe végétal, le nodule, dans lequel les bactéries internalisées (bactéroïdes) fixent l'azote atmosphérique au profit de leur hôte. Les rhizobia ne constituent pas un groupe taxonomique homogène. Ils appartiennent à une quinzaine de genres dispersés au sein des α - and β -protéobactéries. Les rhizobia auraient évolué à partir du transfert horizontal de gènes essentiels à la symbiose, suivi d'une réorganisation du génome d'accueil sous pression de sélection de la plante permettant une activation et/ou optimisation du potentiel symbiotique acquis. Ce scénario évolutif a été reproduit en laboratoire par une approche d'évolution expérimentale. Le plasmide symbiotique du symbiote de *Mimosa pudica*, *Cupriavidus taiwanensis* LMG19424, a été introduit dans la bactérie pathogène de plante *Ralstonia solanacearum* GMI1000. A partir de cette bactérie chimère 18 lignées parallèles ont été évoluées par des cycles successifs d'inoculation à *M. pudica* et ré-isolation des bactéries des nodules. Après 16 cycles d'évolution, trois observations ont été faites : i) les bactéries évoluées ne fixent pas l'azote et l'évolution vers le mutualisme n'est donc pas achevée à ce stade, ii) un gène de fonction inconnue semble important pour l'infection intracellulaire, et iii) les mutations permettant l'acquisition et/ou l'amélioration de l'infection des cellules du nodule semblent également améliorer la nodulation.

Afin d'identifier les conditions favorables à l'émergence du mutualisme dans l'expérience d'évolution et potentiellement dans la nature, nous avons analysé la dynamique spatio-temporelle de deux sous-populations quasi isogéniques de *C. taiwanensis*, l'une fixatrice d'azote (Fix+) et l'autre non fixatrice (Fix-), au cours du processus symbiotique avec *M. pudica*. Nous avons observé une dégénérescence précoce et sélective des Fix-, y compris lorsqu'ils partagent un même nodule avec des Fix+, et établit la cinétique d'expansion des Fix+ au cours du temps. A partir d'un modèle mathématique et de validations expérimentales, nous avons prédit que de rares Fix+ envahiraient une population majoritairement Fix- au cours de cycles successifs de nodulation avec une probabilité fonction de la taille initiale de l'inoculum, du nombre de plantes inoculées et de la longueur des cycles.

Par la suite nous avons étudié le rôle d'un gène du plasmide symbiotique de *C. taiwanensis*, dont la délétion dans l'une des lignées était responsable d'un défaut d'infection intracellulaire. Nous avons montré que ce gène, appelé *noeM*, est un gène de nodulation impliqué dans la biosynthèse de facteurs Nod atypiques où le sucre réducteur est ouvert et oxydé. *noeM* est principalement détecté dans des isolats de plantes appartenant à la tribu des Mimoseae, et particulièrement chez les souches capables de noduler *M. pudica*. Les gènes *noeM* forment un clade phylogénétique à part et spécifique des rhizobia. Un mutant $\Delta noeM$ de *C. taiwanensis* s'est avéré affecté pour la nodulation de *M. pudica*, confirmant son rôle dans la symbiose avec cette légumineuse.

Enfin, l'analyse cytologique détaillée de l'infection racinaire de *M. pudica* par *C. taiwanensis* et quelques souches de *R. solanacearum* portant une mutation adaptative de l'infection intracellulaire a été initiée, afin d'analyser l'impact de ces mutations sur les étapes symbiotiques précoces.

Mots clefs : symbiose, évolution, fixation d'azote, *R. solanacearum*, *C. taiwanensis*, *M. pudica*, infection, sanctions, facteurs Nod

Abstract :

The symbiosis between legumes and bacteria, known as rhizobia, is a complex process resulting in the formation of a novel plant organ, the nodule, in which internalized bacteria (bacteroids) fix nitrogen to the benefit of the host plant. Rhizobia do not form a homogeneous taxonomic group. They belong to a dozen of genera scattered within α - and β -proteobacteria. Rhizobia may have evolved from horizontal transfer of key symbiotic genes, followed by genome remodeling under plant selection pressure, allowing the activation and/or optimization of the acquired symbiotic potential. This evolutionary scenario is being replayed in the laboratory using an experimental evolution approach. The symbiotic plasmid of the *Mimosa pudica* symbiont, *Cupriavidus taiwanensis* LMG19424, was introduced into the plant pathogen *Ralstonia solanacearum* GM11000. 18 parallel lineages were derived from this chimeric ancestor using serial cycles of inoculation with *M. pudica* and re-isolation of bacteria from the nodules. After 16 cycles of evolution, three observations were done: i) the evolved bacteria do not fix nitrogen and evolution towards mutualism is not completed, ii) a gene of unknown function seems to be involved in intracellular infection and iii) the mutations that allow and/or improve intracellular infection also improve nodulation capacity.

To determine conditions that favor the emergence of mutualism in the laboratory and possibly in nature, we analyzed the spatio-temporal dynamics of two quasi-isogenic sub-populations of *C. taiwanensis*, one nitrogen-fixing (Fix+) and the other not (Fix-), along their symbiotic process with *M. pudica*. We observed an early degenerescence of Fix- bacteroids, even when they share a nodule with Fix+, and established the kinetics of Fix+ expansion along time. Using mathematical modeling and experimental validations, we predicted that rare Fix+ will invade a population dominated by non-fixing bacteria during serial nodulation cycles with a probability that is function of initial inoculum, plant population size and nodulation cycle length.

Then, we studied the role of a *C. taiwanensis* symbiotic plasmid gene, whose deletion in one lineage was responsible of intracellular infection defect. We showed that this gene, called *noeM*, is a novel nodulation gene involved in the biosynthesis of atypical Nod factors where the reducing sugar is open and oxidized. *noeM* was mostly found in isolates of the *Mimoseae* tribe, especially in all strains able to nodulate *M. pudica*. The *noeM* genes form a separate phylogenetic clade containing only rhizobial genes. A *noeM* deletion mutant of *C. taiwanensis* was affected for the nodulation of *M. pudica* confirming the role of *noeM* in the symbiosis with this legume.

Last, we initiated the detailed cytological analysis of *M. pudica* root infection by *C. taiwanensis* and a few strains bearing adaptive mutations for intracellular infection, in order to analyze the effect of these mutations on early symbiotic stages.

Key words : symbiosis, evolution, nitrogen fixation, *R. solanacearum*, *C. taiwanensis*, *M. pudica*, infection, sanctions, Nod factors

Liste des abréviations :

3-OH MAME : methyl 3-hydroxymyristate

ADN : acide désoxyribonucléique

ADP : adénosine diphosphate

ARA : acetylene reduction assay

ARN : acide ribonucléique

ATP : adénosine triphosphate

CCaMK : calcium and calmodulin-dependent kinase

CCSP : voie de signalisation symbiotique commune

dpi: jours post inoculation

EE : expérience d'évolution

EPS : exopolysaccharides

FN(s) : facteur Nod

GFP : green fluorescence protein

HR : Hypersensitive response

Hrc : *hrp* conserved

Hrp : hypersensitive response and pathogenicity

IRLC : inverted repeated legume clade

KPS : polysaccharide capsulaire

LPS : lipopolysaccharide

LRR : leucine rich repeat

LysM : lysine motif

N₂ : diazote

O₂ : dioxygène

Pb : paire de bases

PCR : polymerase chain reaction

Phc : phenotypic conversion

Prh : plant regulator of *hrp* genes

sp. : species

SST3 : système de sécrétion de type 3

Table des matières

CHAPITRE I : INTRODUCTION	1
I.1 LA FIXATION BIOLOGIQUE DE L'AZOTE	1
I.1.1 <i>Le complexe nitrogénase</i>	1
I.1.2 <i>Les bactéries fixatrices d'azotes</i>	2
I.2 LA SYMBIOSE RHIZOBIUM-LEGUMINEUSE	5
I.2.1 <i>Les acteurs de la symbiose fixatrice d'azote</i>	5
I.2.1.1 Les légumineuses	5
I.2.1.2 Les rhizobia	6
I.2.1.3 La spécificité de la symbiose	7
I.2.1.4 Le processus symbiotique	7
I.2.2 <i>Déterminants génétiques de la symbiose rhizobium-légumineuse</i>	10
I.2.2.1 Les stratégies de nodulation et d'infection	11
I.2.2.2 Le dialogue moléculaire de la symbiose FN dépendante.....	11
I.2.2.2.1 Les flavonoïdes.....	12
I.2.2.2.2 Reconnaissance des flavonoïdes et induction des gènes <i>nod</i>	12
I.2.2.2.3 Les facteurs Nod.....	13
I.2.2.3 Autres facteurs d'infection.....	13
I.2.2.3.1 Les polysaccharides de surface.....	13
I.2.2.3.2 Les systèmes de sécrétion.....	15
I.2.2.4 Voies de signalisation végétales conduisant à l'infection et à l'organogénèse du nodule	15
I.2.2.4.1 Perception des NFs.....	16
I.2.2.4.2 Transduction du signal FNs	16
I.2.2.4.3 Libération des rhizobia et différenciation en bactéroïdes.....	17
I.2.3 <i>Évolution de la symbiose fixatrice d'azote</i>	17
I.2.3.1 Les plantes	18
I.2.3.2 Les bactéries	18
I.2.3.2.1 Origine par transfert horizontal de gènes	18
I.2.3.2.2 Étape d'optimisation post-HGT	20
I.2.3.3 Évolution du mutualisme.....	21
I.3. ÉVOLUTION EXPERIMENTALE DE SYMBIOTES DE LEGUMINEUSES	21
I.3.1. <i>L'évolution expérimentale</i>	22
I.3.2. <i>Le modèle Ralstonia solanacearum/Cupriavidus taiwanensis /Mimosa pudica</i>	23
I.3.2.1 Cupriavidus taiwanensis	23
I.3.2.2 Ralstonia solanacearum.....	24
I.3.2.3 Mimosa pudica.....	27
I.3.3. <i>L'expérience d'évolution</i>	28
I.3.3.1. Design de l'expérience.....	28
I.3.3.2. Analyse des mécanismes moléculaires et évolutifs de l'adaptation	28
I.3.4. <i>Projet de thèse</i>	29
CHAPITRE II : LE CONTROLE SPATIO-TEMPOREL DU MUTUALISME CHEZ LES LEGUMINEUSES FAVORISE LA DISSEMINATION DE LA SYMBIOSE FIXATRICE D'AZOTE	31
II.1. INTRODUCTION	31
II.2. ARTICLE	32
CHAPITRE III: IDENTIFICATION D'UN NOUVEAU GENE NOD IMPLIQUE DANS UNE STRUCTURE ATYPIQUE DES FACTEURS NOD	54
III.1. INTRODUCTION	54
III.2. ARTICLE	55
CHAPITRE IV : RESULTATS COMPLEMENTAIRES	70
IV.1. ANALYSES COMPLEMENTAIRES DU GENE <i>noeM</i>	70
IV.1.1 <i>Spécificité fonctionnelle de noeM</i>	70
IV.1.2. <i>Impact de noeM chez R. solanacearum</i>	71
IV.2.1. <i>Analyse préliminaire des étapes précoces de l'interaction avec M. pudica</i>	73
CHAPITRE V : CONCLUSIONS ET DISCUSSION GENERALE	76

V. 1. <i>NOEM</i> , UN NOUVEAU GENE DE SPECIFICITE D'HOTE	76
Quelle est la fonction précise de NoeM ?.....	76
Comment noeM s'est spécialisé dans la symbiose ?.....	78
Comment expliquer la sélection pour Δ noeM dans l'expérience d'évolution ?.....	79
V. 2. LA DISSEMINATION DU CARACTERE FIXATEUR D'AZOTE.....	80
Un contrôle au niveau cellulaire.....	80
Un mécanisme de sanction végétal?.....	81
Évolution naturelle et expérimentale du mutualisme.....	81
V.3. LES ETAPES TARDIVES DE L'INFECTION DU NODULES SONT-ELLES INFLUENCEES DES LES PREMIERES ETAPES DE L'INFECTION DU POIL ABSORBANT ?	82
REFERENCES BIBLIOGRAPHIQUES :	84

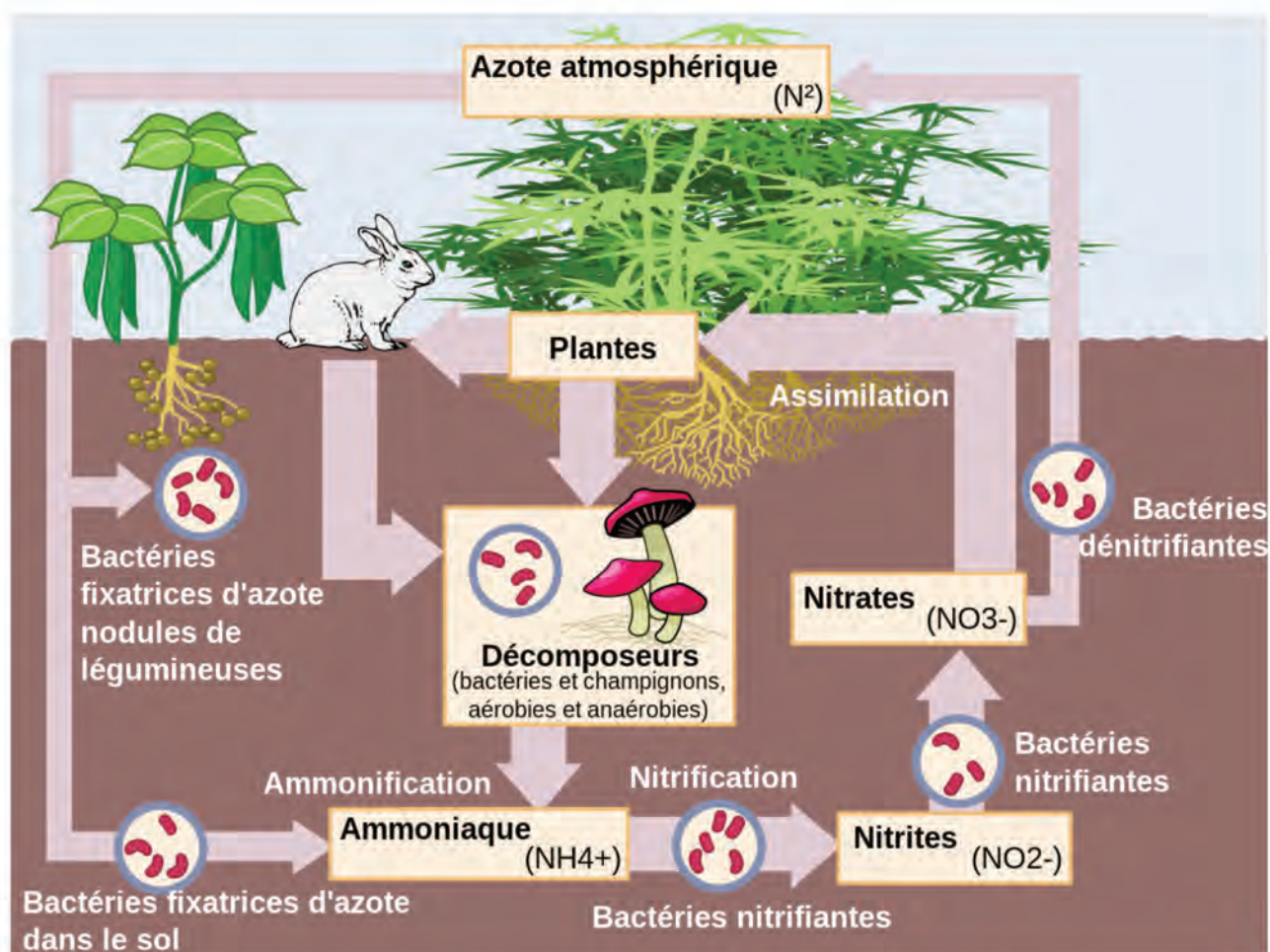


Figure I.1.1.1: Représentation schématique de la fixation biologique de l'azote dans le sol. Les bactéries fixatrices d'azotes du sol réduisent le diazote en ammoniac (NH_4^+) qui est transformé en nitrite (NO_2^-) puis nitrate (NO_3^-) par les bactéries nitrifiantes du sol des milieux aquatiques. Ce cycle biologique se poursuit par l'absorption des nitrates ou de l'ammonium par le système racinaire végétal afin d'être convertis en acides aminés permettant leur transfert aux animaux via les réseaux trophiques. L'azote est également réduit par des bactéries en symbiose avec des plantes qui leur fournissent directement l'azote assimilable. Ensuite, la décomposition des organismes par des bactéries saprophytes permet le retour de l'azote dans le sol sous forme d'ammoniac. Le cycle est bouclé grâce à la réduction d'une part des nitrates présents dans le sol en diazote, par des bactéries dénitrifiantes.

Chapitre I : Introduction

I.1 La fixation biologique de l'azote

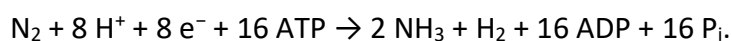
L'azote est l'élément majoritaire de l'air atmosphérique (l'azote libre représente près de 80% de l'air sous forme de diazote (N₂)) et est aussi l'un des constituants majeurs du vivant puisqu'on le retrouve dans les acides nucléiques et les acides aminés qui possèdent tous un groupe amine (-NH₂ ou -NH-). Présent en quantité inépuisable, l'azote atmosphérique n'est assimilable que par certains procaryotes qui ont la capacité de transformer le N₂ atmosphérique en azote assimilable. Ce processus, s'appelle la fixation biologique de l'azote.

Afin de répondre à leurs besoins, les eucaryotes doivent assimiler l'azote combiné (qui comprend l'azote organique et l'azote minéral). L'azote combiné représente 0,001% de l'azote total de la biosphère ce qui fait de cet élément un facteur de croissance limitant (Newton 2000).

Au début du XX^{ème} siècle Fritz Haber et Carl Bosh ont développé une méthode permettant de fixer l'azote industriellement, en faisant réagir l'azote atmosphérique avec des métaux à températures élevées et sous pression. Ce processus permettant ainsi la production d'engrais est cependant coûteux et polluant (Erisman et al. 2008). L'azote peut aussi être fixé par des mécanismes naturels non biologiques, c'est le cas par exemple de la foudre qui va transformer l'azote et l'oxygène en oxyde d'azote (Dawson 1980).

I.1.1 Le complexe nitrogénase

La conversion de l'azote en ammoniac par les procaryotes représente la majorité de la fixation biologique annuelle (**Figure I.1.1.1**) (Orr et al. 2011) et est rendue possible par un complexe enzymatique appelé nitrogénase dont le bilan de réaction est le suivant :



Trois sortes de nitrogénases ont été identifiées. Les différences qui caractérisent les complexes enzymatiques se situent au niveau de la nature des atomes composant leur hème métallique. On retrouve ainsi, la V-nitrogénase qui possède des atomes de vanadium (Rehder 2000; Lee et al. 2010), la Fe-nitrogénase possédant du fer (Schneider et al. 1997) et enfin la Mo-nitrogénase. Cette dernière est la plus étudiée puisque c'est celle que l'on retrouve le plus communément chez les diazotrophes et chez tous les symbiotes de légumineuses. Elle se compose de deux métalloprotéines dont la première est une dinitrogénase réductase de type protéine-Fer. C'est un dimère de 60 KDa constitué de deux sous unités identiques. Chacune de ces deux sous-unités présentent un site de fixation aux molécules de Mg²⁺-ATP. La seconde métalloprotéine est une dinitrogénase de type protéine MoFe en forme de tétramère d'environ 220KDa constitué de deux sous unités α et de deux sous unités β . A cela s'ajoutent

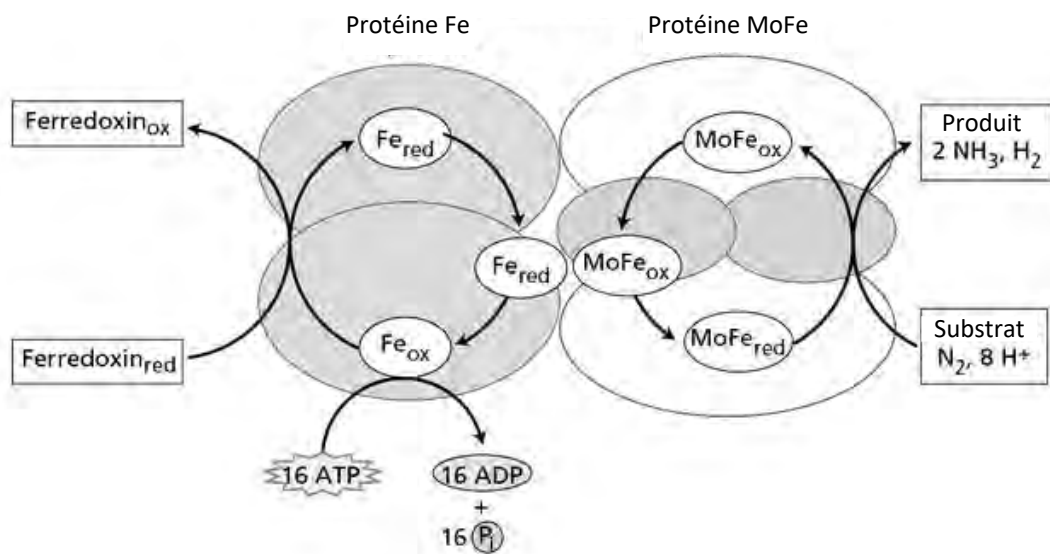


Figure I.1.1.2: Schéma du complexe Mo-nitrogénase. Le schéma représente la structure et l'opération de réduction de l'azote par le complexe enzymatique. Les électrons sont transférés de la ferredoxin (ou flavodoxin) via l'azoferredoxin (protéine Fe) vers la molybdoferredoxin (protéine MoFe). Chaque mol d'azote fixée nécessite 16 mol d'ATP hydrolysée par la protéine nifH. Adapté de (Rascio et al. 2013).

deux atomes de molybdène présents dans deux cofacteurs polypeptidiques riches en atome de Fer (FeMo-CO) et impliqués dans le processus de réduction (**Figure I.1.1.2**) (Burgess and Lowe 1996; Seefeldt et al. 2009; Rees and Howard 2000). Ces complexes enzymatiques sont codés par un ensemble de gènes spécifiques appelés *vnf* pour la nitrogénase à vanadium, *anf* pour celle qui comporte uniquement du fer et gènes *nif* pour la nitrogénase à molybdène.

Chez *Klebsiella pneumoniae*, 20 gènes *nif* sont nécessaires pour la synthèse du complexe enzymatique de la nitrogénase à molybdène, ainsi qu'à son bon fonctionnement. Chez les cyanobactéries, 16 gènes *nif* sont présents alors que le nombre de gènes *nif* impliqués dans la mise en place de la fixation est très variable chez les rhizobia (Latysheva et al. 2012). Ainsi, *S. meliloti* et *R. leguminosarum* *bv. viciae* possèdent le plus petit cluster de gènes *nif* avec, respectivement, 9 et 8 gènes *nif* dont les gènes *nifABDEHKN* qui sont commun à tous les microorganismes fixateurs.

Les gènes *nifHDK* codent pour la structure du complexe nitrogénase où *nifH* code pour la dinitrogénase reductase et *nifDK* encodent les sous-unité α et β de la dinitrogénase (Kneip et al. 2007). La plupart des autres gènes, comme *nifENUSVWXBQ*, sont impliqués dans l'assemblage et l'incorporation des atomes de fer et du molybdène dans les sous unité de la nitrogénase. Les gènes *nifF* et *nifJ* encodent les protéines impliquées dans les transferts d'électrons qui prennent place dans le processus de réduction chez *Klebsiella* (Dixon and Kahn 2004). NifF et NifJ étant absents chez les rhizobia, ce sont probablement les gènes *fixABCX* qui remplacent *nifF* et *nifJ* (Masson-Boivin et al. 2009). De nombreux autres gènes *nif* sont absents chez les rhizobia, comme *nifSUX*, qui sont probablement remplacés par les gènes paralogues et de ménage *icsS* et *icsA* (Masson-Boivin et al. 2009). Chez les fixateurs symbiotiques, la régulation des gènes *nif* se fait, entre autres, via le gène de régulation *nifA* dont la synthèse et l'activité sont régulés par l'oxygène (Dixon and Kahn 2004; Masson-Boivin et al. 2009).

La sensibilité de la nitrogénase à l'oxygène nécessite des stratégies d'adaptation physiologique afin d'en protéger l'enzyme. Ainsi, certains microorganismes vont augmenter excessivement leur respiration, produire des barrières empêchant la diffusion de l'oxygène ou encore compartimenter la nitrogénase (Dixon and Kahn 2004). La symbiose rhizobium-légumineuse est un bon exemple de stratégie d'adaptation dans laquelle le cortex nodulaire fait office de barrière de diffusion à l'oxygène (Dixon and Kahn 2004). Dans les nodules, la plante hôte va synthétiser une protéine, appelée légghémoglobine, qui permet à la bactérie de respirer tout en protégeant le complexe enzymatique des effets de l'oxygène (Appleby 1984). Ainsi, d'importants changements de régulations sont observés chez les fixateurs symbiotiques dans leur transition de vie libre à vie symbiotique (Dixon and Kahn 2004).

I.1.2 Les bactéries fixatrices d'azotes

Parmi les organismes vivants, la fixation du N₂ atmosphérique est une caractéristique exclusive des Archées et des Eubactéries. On dénombre chez les Archées deux ordres contenant des procaryotes fixateurs d'azote qui sont les *Methanosarcinales* et les

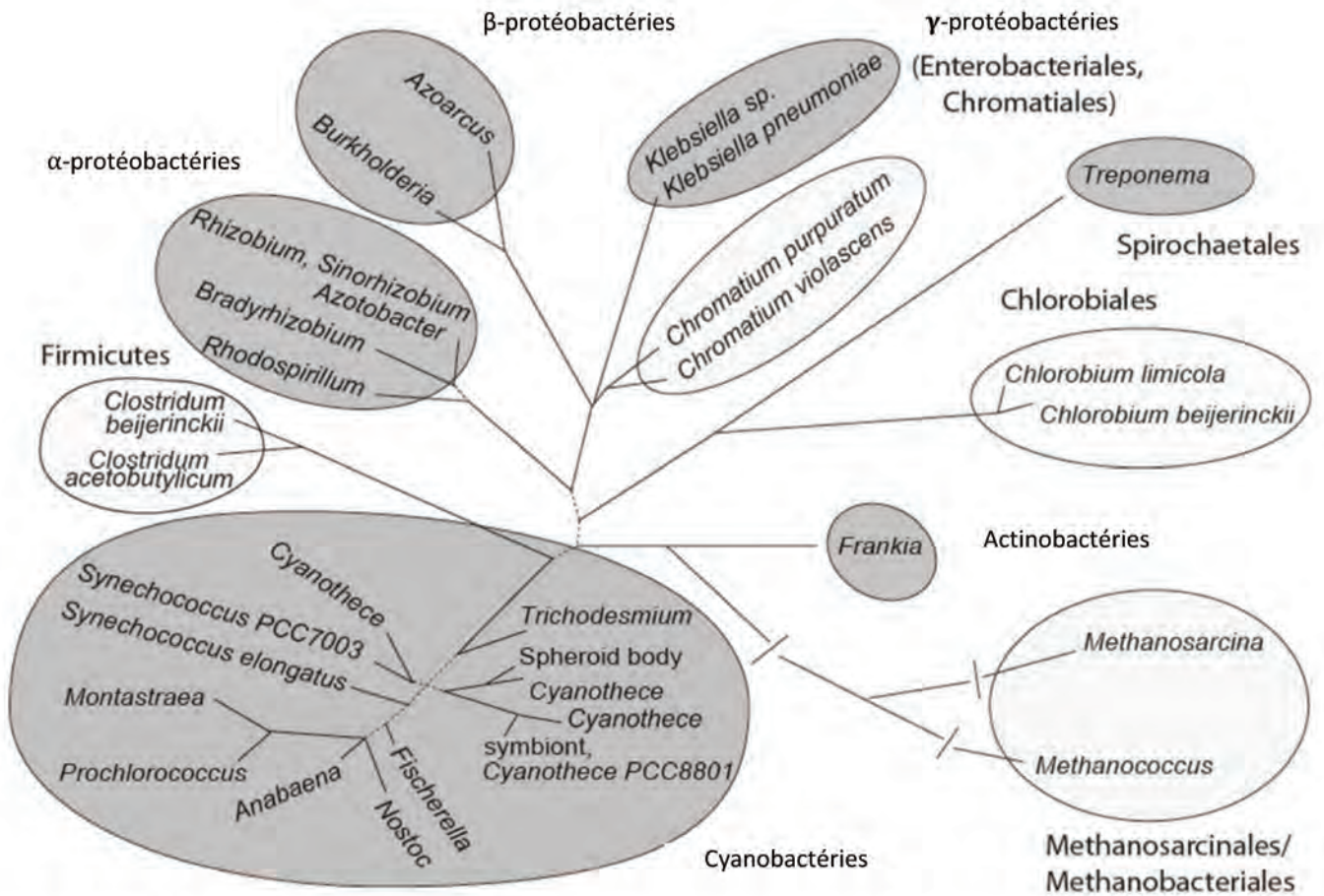


Figure I.1.2.1: Arbre phylogénétique des procaryotes symbiotiques ou non capable de fixer l'azote. Les différentes divisions de bactéries fixatrices d'azote (classes/ordres pour les protéobactéries) sont indiquées. Les groupes contenant les espèces symbiotiques sont colorés en gris et les non-symbiotiques sont marqués par les ellipses blanches. Adapté de (Kneip et al. 2007).

Methanobacteriales (Kneip et al. 2007). Parmi les Eubactéries fixatrices d'azote on retrouve des espèces faisant partie des embranchements des Cyanobactéries et Chlorobactéries, ainsi que des espèces de bactéries Gram positive du groupe des Actinobactéries et des Firmicutes, mais aussi des bactéries Gram-négatives du groupe des Protéobactéries et enfin des espèces des ordres des *Chlorobiales* et *Spirochétales* (Kneip et al. 2007) (**Figure I.1.2.1**).

Les cyanobactéries sont un clade diversifié tant biochimiquement que morphologiquement jouant un rôle majeur dans l'environnement de par leur productivité primaire mais aussi au niveau économique avec leur rôle potentiel dans les biocarburants ou leur utilisation comme engrais vert (Latysheva et al. 2012). Elles existent probablement depuis plus de 2.9 milliards d'années et de nombreuses espèces ont été identifiées comme étant capables de fixer l'azote. Parmi les Actinobactéries, il a été démontré que de nombreuses espèces d'*Arthrobacter*, *Agromyces*, *Corynebacterium*, *Mycobacterium*, *Micromonospora*, *Propionibacteria*, *Streptomyces* mais aussi des *Frankia* sont capables de fixer le N₂ (Sellstedt and Richau 2013). Les protéobactéries fixatrices comptent de multiples genres capables de fixer l'azote atmosphérique dont les plus étudiés sont *Klebsiella pneumoniae* ainsi que les α - et β -rhizobia.

Ces organismes évoluent dans la plupart des écosystèmes terrestres et marins mais aussi dans des environnements aux conditions extrêmes (Masson-Boivin and Sachs 2018; Gaby et al. 2018). La fixation biologique de l'azote se produit via deux modes de vie bien distincts : en vie libre, que l'on retrouve chez les diazotrophes, ou en symbiose.

Parmi les fixateurs libres ont été identifiés des Cyanobactéries, des Protéobactéries, des Firmicutes et des Archées. Ils se caractérisent par leur mode de vie anaérobie, anaérobie facultative ou aérobie obligatoire (Zhan and Sun 2012; Kneip et al. 2007) :

- Les fixateurs anaérobies, comme *Clostridium* ou *Citrobacter*, ne peuvent tolérer l'oxygène même lorsqu'il ne fixent pas l'azote. On les retrouve dans des habitats pauvres en oxygène comme le sol ou les matières végétales en décomposition. On peut retrouver certaines archées fixatrices, comme *Methanococcus*, dans les boues ou les intestins d'animaux (French et al. 1976; Minamisawa et al. 2004; Dixon and Kahn 2004).
- Il existe quelques espèces fixatrices connues capable de vivre de façon anaérobie ou aérobie, *Klebsiella pneumoniae*, *Paenibacillus polymyxa* ou encore *Bacillus macerans*. Néanmoins, afin de pouvoir fixer l'azote des conditions anaérobies sont requises (Tilden and Hudson 1942; Lal and Tabacchioni 2009; Dixon and Kahn 2004).
- Les aérobies obligatoires ne peuvent se développer sans oxygène. La bactérie Gram-négative *Azotobacter vinelandii* est la plus étudiée des diazotrophes aérobies. Cette dernière augmente fortement sa respiration afin de réduire la quantité d'oxygène et empêcher tout dommage de sa nitrogénase (Kneip et al. 2007). La capacité photosynthétique des cyanobactéries complique leur habilité à fixer l'azote. Afin de contrebalancer ce problème, les cyanobactéries utilisent deux mécanismes spatio-temporels afin de séparer ces activités. L'utilisation d'un cycle circadien biologique



Figure 1.1.2.2: Les symbioses plante-cyanobactéries représentatives. (a) Photomicrographie d'une espèce de *Rhizosolenia* contenant *Richelia intracellularis* (flèche). Bar, 30 μ m. (b) Le bipartite cyanolichen *Peltigera canina*. Le cyanobiotte est localisé dans le thalle sombre du foliose. Bar, 2.5 cm. (c) Le tripartite cyanolichen *Peltigera aphthosa*. Bar, 2.5 cm. A noter le *Nostoc* contenant en externe les cephalodia (points noirs) sur le thalle vert, qui contient l'algue *Cccomyxa* (d) Le thalle gametophyte de l'espèce *Anthoceros*. Le *Nostoc* cyanobiotte est situé in dans les cavités (flèche). Bar, 3 mm. (e) Le sporophyte de l'espèce *Azolla*. La feuille dorsal (flèche) contient des cavités dans lesquelles les cyanobiontes sont situés. Bar, 1 mm. (f) Le cycadophyte de l'espèce *Bowenia* avec des racines coralloïdes (flèche). Bar, 2.5 cm. (g) Une coupe de racine coralloïde. Le cyanobiotte *Nostoc* est situé dans la zone mid-cortical de la racine, vue ici comme le cercle bleuté (flèche). Bar, 1 mm. (h) *Gunnera magellanica*. L'accolade indique la portion sectionnée en (i). Bar, 1.3 cm. (i) Une coupe longitudinale de tige apicale *G. magellanica* montrant les glandes qui contiennent (flèches) *Nostoc* cyanobiotte. Bar, 4 mm. Adapté de (Rai et al. 2000)

alternant photosynthèse et fixation d'azote sépare temporellement ces activités et une différenciation cellulaire en cellule spécialisée appelée hétérocyste, les dissocie spatialement (Kumar et al. 2010).

Il existe trois catégories de microorganismes connus pour être capables de fixer le N₂ en symbiose avec des organismes eucaryotes : les cyanobactéries, les *Frankia* et des protéobactéries de genres divers communément appelés rhizobia.

Un certain nombre de cyanobactéries sont capables de symbioses mutualistes avec des protistes, des animaux ou des végétaux. Parmi les végétaux, on retrouve des champignons, des diatomées, des bryophytes mais aussi des ptéridophytes, des gymnospermes et des angiospermes (Rai et al. 2000; Kneip et al. 2007). Les interactions mutualistes impliquant les cyanobactéries hétérocystes avec les plantes sont des symbioses fixatrices d'azotes sont spécifique, une espèce d'hôte ne s'associant qu'avec un genre particulier de cyanobactérie (Rai et al. 2000). Les conditions naturelles requises à la mise en place de ce mutualisme sont encore inconnues. Les cyanobactéries peuvent être présentes dans la plante de façon extracellulaire ou intracellulaire, facilitant les échanges entre les partenaires (Rai et al. 2000). La localisation du symbiote est variable en fonction de son hôte, l'on peut ainsi retrouver des cyanobactéries symbiotiques, entre autres, dans des cavités foliaires, de simples cellules ou des cavités de thalles (**Figure I.1.2.2**).

Les bactéries du genre *Frankia* et les rhizobia forment une symbiose dite « nodulaire ». Ces bactéries induisent la formation d'organes spécifiques, les nodules ou nodosités, sur la racine (et quelquefois la tige) de leurs hôtes. Ces hôtes sont restreints aux plantes appartenant à quelques ordres des Angiospermes Dicotylédones.

Les *Frankia* sont des bactéries actinomycètes filamenteuses et telluriques capables d'interaction symbiotique avec les plantes appartenant aux Rosales, Fagales et Curcubitales communément appelées plantes actinorhiziennes (Masson-Boivin et al. 2009). Les souches de *Frankia* peuvent être divisée en 4 clusters, dont 3 d'entre eux (les clusters I, II, III) contiennent des souches symbiotiques (Nguyen et al. 2016). Cette symbiose se met en place lorsque la plante produit un signal composé de flavonoïdes et d'isoflavonoïdes auquel les *Frankia* répondent en produisant un signal permettant l'activation du processus symbiotique de la plante hôte en activant la voie de signalisation commune à la symbiose décrite chez les légumineuses. Le signal produit par les *Frankia* spp. n'a à ce jour pas encore été clairement identifié (Oldroyd 2013; Sellstedt and Richau 2013).

Les rhizobia sont des bactéries du sol Gram négative appartenant au phylum des protéobactéries et capables d'entrer en symbiose avec les plantes de la famille des légumineuses (Fabales), à l'exception de *Parasponia* (Rosales). Cette interaction est décrite en détail dans les paragraphes suivants.

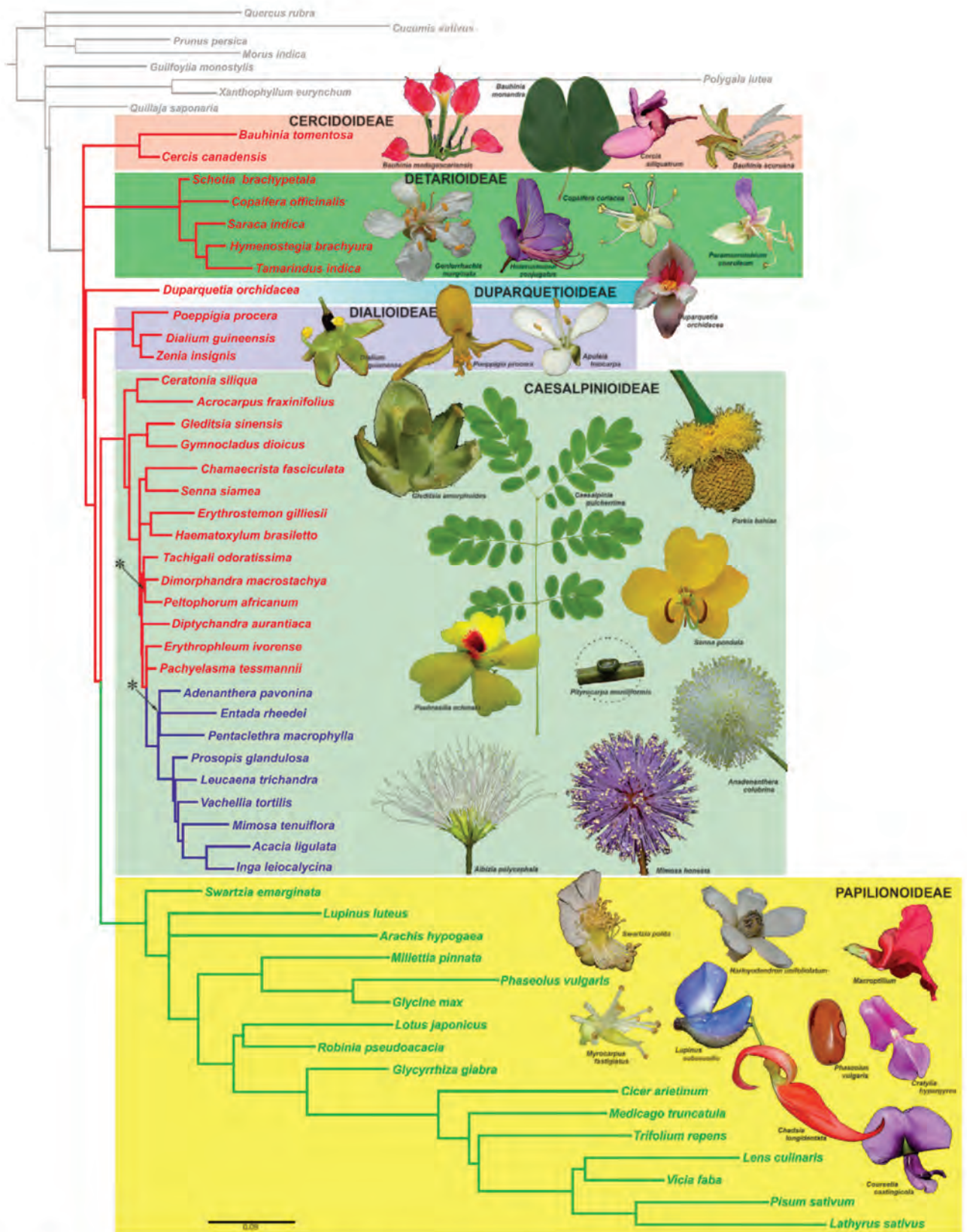


Figure I.2.1.1.1: Phylogénie et classification des sous familles des Leguminosae. Les six sous familles sont indiquées par les encadrements colorés à droite de la phylogénie. Les branches colorées indiquent l'ancienne classification de sous famille des Leguminosae: en rouge les anciens Caesalpinioideae, en bleu les Mimosoideae et en vert les Papilionoideae (Azani et al. 2017)

I.2 La symbiose rhizobium-légumineuse

La symbiose qui s'établit entre les rhizobias et les légumineuses est l'interaction mutualiste la plus étudiée, et ce depuis plus de 100 ans. La possibilité de manipuler génétiquement ces bactéries ainsi que l'intérêt agronomique de nombreuses légumineuses (pois, sojas etc.), explique l'intérêt scientifique pour cette symbiose. La symbiose rhizobium-légumineuse est une interaction mutualiste très spécifique.

I.2.1 Les acteurs de la symbiose fixatrice d'azote

I.2.1.1 Les légumineuses

Avec plus de 19500 espèces recensées, les légumineuses forment la troisième plus grande famille de plantes réparties en 765 genres différents (Wang et al. 2017; Azani et al. 2017). Les légumineuses sont présentes sur tous les continents et sont considérées comme étant, après les céréales, la seconde famille de plantes la plus importante économiquement. Anciennement divisée en trois sous-familles (*Papilionaceae*, *Mimosaceae* et *Caesalpinaceae*), les Fabacées ont été reclassées et se subdivisent maintenant en 6 sous-familles comprenant les *Cercidoideae*, les *Detarioideae*, les *Duparquetioideae*, les *Dialioideae* et enfin les *Caesalpinionideae* et *Faboideae* chez qui l'on retrouve les espèces capables d'établir une association symbiotique fixatrice d'azote avec les rhizobias ou *Frankia* (**Figure I.2.1.1.1**) (Azani et al. 2017).

La sous-famille des *Caesalpinionideae* est très hétérogène avec 148 genres (incluant ceux de l'ancienne sous-famille des *Mimosoideae*) qui se divisent en environ 4400 espèces, essentiellement des arbustes, des lianes évoluant dans des régions tropicales et sub-tropicales mais aussi quelques arbres de régions tempérées (Azani et al. 2017).

70% des espèces de légumineuses (96% des *Papilionaceae*, 90% des *Mimosaceae* et 22% des *Caesalpinaceae*) sont capables de nodulation avec les rhizobias (Sprent 2009).

La sous-famille des *Faboideae*, aussi connue en tant que *Papilionoideae*, est la sous-famille la plus large des Fabacées avec plus de 14000 espèces réparties dans 503 genres différents. Elle est d'une grande importance économique puisqu'elle comprend de nombreuses espèces cultivées pour l'alimentation humaine, animale ou encore ornementale comme le soja, la luzerne, haricot ou encore le trèfle (Azani et al. 2017).

De par leur intérêt agronomique les légumineuses ont fait rapidement partie des plantes les mieux caractérisées génétiquement. Ainsi, deux légumineuses ont été utilisées très tôt comme plantes modèles dans l'étude de la symbiose fixatrice d'azote : *Medicago truncatula* (*M. truncatula*) et *Lotus japonicus* (*L. japonicus*) (Cook 1999). Le petit génome (500 Mb) séquencé diploïde ($2n=16$) de *M. truncatula*, le nombre élevé de graines, son temps de génération court ainsi que ses nombreux écotypes auto-fertilisables ont fait de cette plante

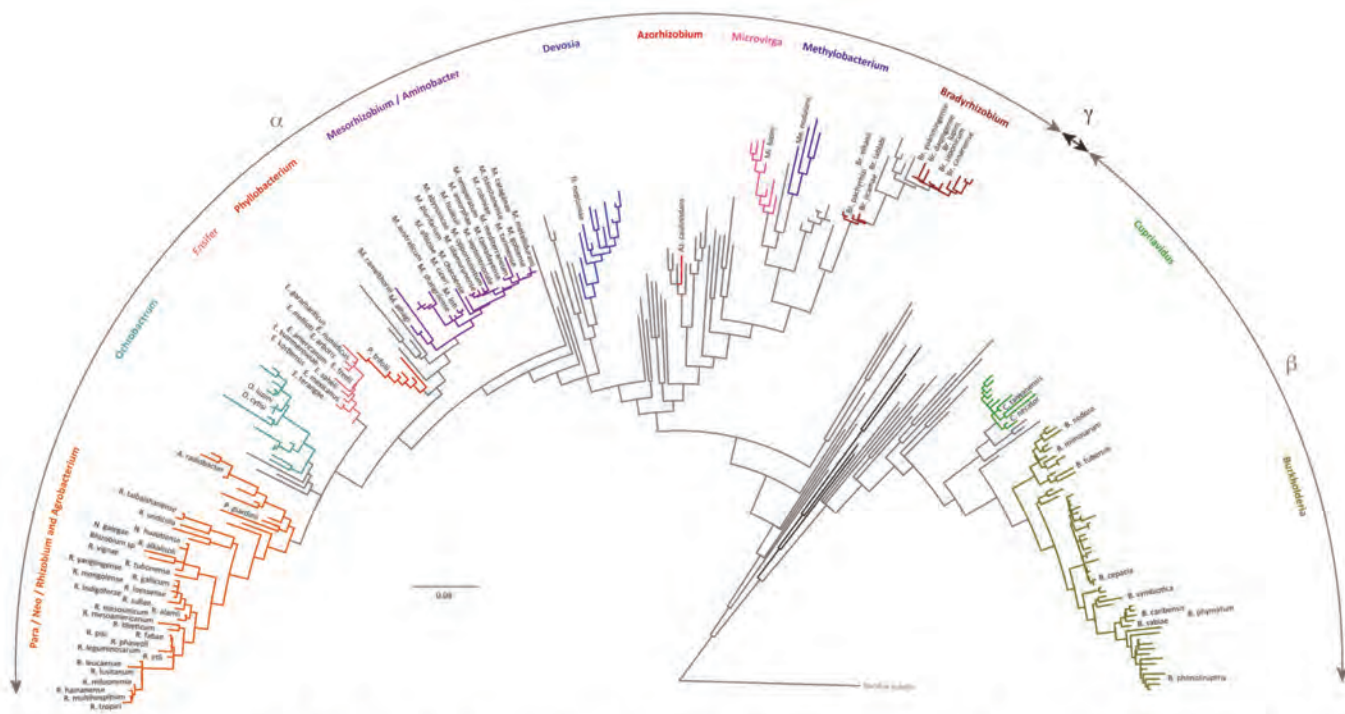


Figure I.2.1.2.1: Phylogénie d'une sélection de protéobactérie montrant la vaste distribution des Rhizobia. Arbre construit à partir de séquences d'ARNr 16S de souches obtenues sur RDP 11.4 (<http://rdp.cme.msu.edu/>) par Mega 6.06 en utilisant 1431 alignements et un maximum de vraisemblance basé sur le modèle Tamura-Nei. Les taxons /genres actuellement connus posséder des rhizobiums sont indiqués en couleur et les espèces rhizobiennes sont indiquées. (Remigi et al. 2016).

un modèle très utilisé (Cook 1999; Young et al. 2011). Il en est de même pour *L. japonicus* qui possède un court cycle de vie (environ 3 mois), une capacité d'autofécondation et un petit génome (472 Mb) diploïde ($2n=12$) qui a été séquencé et assemblé ainsi que de nombreux mutants disponibles (Sato et al. 2008). Par la suite, d'autres travaux de génétique ont été réalisés sur différentes espèces de légumineuses afin de mieux comprendre les interactions rhizobium-légumineuse, c'est le cas par exemple de *Pisum sativum* ou *Medicago sativa* (Kaló et al. 2004; Blair et al. 2003).

I.2.1.2 Les rhizobia

Le terme rhizobia (singulier, rhizobium) est un terme fonctionnel désignant les bactéries capables de former des nodosités avec des légumineuses (Remigi et al. 2016). Il regroupe des bactéries aérobies du sol appartenant à une centaine d'espèces réparties dans 18 genres d' α - et β -protéobactéries (**Figure I.2.1.2.1**) (Tang & Capela, 2019). Ces bactéries sont phylogénétiquement mais également métaboliquement et génétiquement très diversifiées. La classification des rhizobia est en perpétuelle évolution avec de nouvelles espèces et de nouveaux genres découverts régulièrement. Pendant longtemps, les seules bactéries capables de fixer l'azote en symbiose avec des légumineuses appartenaient à quelques genres bactériens de la sous-classe des α -Protéobactéries. Ce n'est qu'en 2001, que de nouveaux genres appartenant aux β -Protéobactéries ont été découverts faisant évoluer considérablement la notion de rhizobium et le caractère fonctionnel de ce terme (Chen et al. 2001; Moulin et al. 2001; Rivas et al. 2002). Toutes les espèces de rhizobia issues des genres d' α -protéobactéries, communément appelées α -rhizobia, appartiennent à l'ordre des Rhizobiales tandis que celles des genres β -protéobactéries (les β -rhizobia) appartiennent à l'ordre des Burkholderiales. Parmi les α -rhizobia les plus connus on retrouve les genres *Rhizobium*, *Bradyrhizobium*, *Ensifer* (anciennement appelé *Sinorhizobium*) mais aussi *Azorhizobium*, *Methylobacterium* ou encore *Mesorhizobium*. L'ordre des Burkholderiales contient les genres symbiotiques *Burkholderia* et *Cupriavidus*, symbiotes principaux des plantes du clade des *Mimosae* (Gyaneshwar et al. 2011). Il est à noter qu'un même genre ou une même espèce peut contenir à la fois des souches symbiotiques et des souches non symbiotiques (Masson-Boivin et al. 2009). Par exemple *Cupriavidus taiwanensis* contient des souches isolées de nodosité et des souches isolées de patients atteints de mucoviscidose (Chen et al. 2001).

En accord avec leur diversité phylogénétique, les rhizobia possèdent de nombreuses propriétés métaboliques différentes (Masson-Boivin et al. 2009). Par exemple, *Azorhizobium caulinodans*, symbiote du genre *Sesbania*, est le seul rhizobium connu pour être un vrai diazotrophe, capable de fixer l'azote en vie libre (Dreyfus et al. 1988). *Methylobacterium nodulans* est une bactérie méthylotrophe, c'est-à-dire capable d'utiliser le méthanol (Sy et al. 2001). Certains *Bradyrhizobium* symbiotes des légumineuses du genre *Aeschynomene* sont photosynthétiques (Giraud and Fleischman 2004).

I.2.1.3 La spécificité de la symbiose

L'une des caractéristiques de la symbiose rhizobium-légumineuse est la spécificité qui existe entre les deux partenaires (Schultze et al. 1991; Debelle et al. 1992). Un rhizobium ne peut établir une symbiose qu'avec un certain nombre, plus ou moins important, d'espèces végétales, défini par son spectre d'hôte. Par exemple, certains rhizobia comme *Sinorhizobium* sp. NGR234 et *Sinorhizobium fredii* USDA257 sont des symbiotes à large spectre capables de noduler de nombreuses plantes (près de 120 genres, dont la non légumineuse *Parasponia* pour NGR234 et plus de 77 genres pour USDA257) (Pueppke and Broughton 1999). Au contraire, d'autres interactions sont très spécifiques, c'est le cas notamment d'*Azorhizobium caulinodans* qui ne peut s'associer qu'avec les plantes du genre *Sesbania* (Dreyfus et al. 1988). Il existe peu de rhizobia symbiotes d'une seule espèce/genre de légumineuse (Perret et al. 2000). La majorité des souches rhizobiennes présentent un spectre d'hôte à spécificité modérée et peuvent noduler plusieurs plantes. C'est le cas de *Sinorhizobium meliloti* qui est capable de noduler les genres *Medicago*, *Melilotus* et *Trigonella* (Krishnan and Pueppke 1991).

De la même façon, une espèce de légumineuse pourra mettre en place une association symbiotique avec une ou plusieurs espèces de rhizobia, d'un ou de plusieurs genres différents, comme par exemple *Mimosa pudica* qui est nodulée par des α - et des β -rhizobia (Melkonian et al. 2014). D'autres légumineuses comme *Vigna unguiculata* (Lewin et al. 1987), *Phaseolus vulgaris*, mais aussi les plantes des genres *Glycine* (Martínez-Romero 1994), *Acacia* et *Sesbania* (De Lajudie et al. 1994; Nick et al. 1999) sont capables de s'associer avec de multiples espèces de rhizobia. Au contraire, certaines légumineuses ont des spectres d'hôtes très restreint, comme *Galega* qui ne peut noduler qu'avec *Rhizobium galegae* (Lindström 1989; Terefework et al. 1998). Certaines plantes, comme le pois afghan nécessitent des dispositions particulières pour pouvoir mettre en place l'association symbiotique. Cette plante présente une résistance à la nodulation et n'est nodulée que par certaines souches de rhizobia possédant un gène *nod* bien particulier (*nodX*) (Firmin et al. 1993).

I.2.1.4 Le processus symbiotique

La mise en place de la symbiose fixatrice d'azote entre le symbiote et sa plante hôte est un processus complexe qui inclut différentes étapes plus ou moins chevauchantes : l'infection des racines, la formation des nodules et l'infection des cellules du nodule suivi de la différenciation des bactéries intracellulaires en bactéroïdes et la fixation de l'azote. Il y a une grande variabilité dans la symbiose rhizobium-légumineuse. En particulier, la localisation des nodules, leur forme et leur anatomie ainsi que le mode d'infection et le niveau de différenciation en bactéroïdes varient selon les systèmes (Masson-Boivin et al. 2009).

Le mode d'infection n'est pas le même pour tous les rhizobia et est classé en différentes catégories selon la façon dont les bactéries franchissent l'épiderme et progressent dans le cortex et si elles induisent la formation de cordons d'infection ou non. Le mode d'infection le plus connu et le plus étudié est celui dit « root hair » puisque c'est celui que l'on retrouve chez

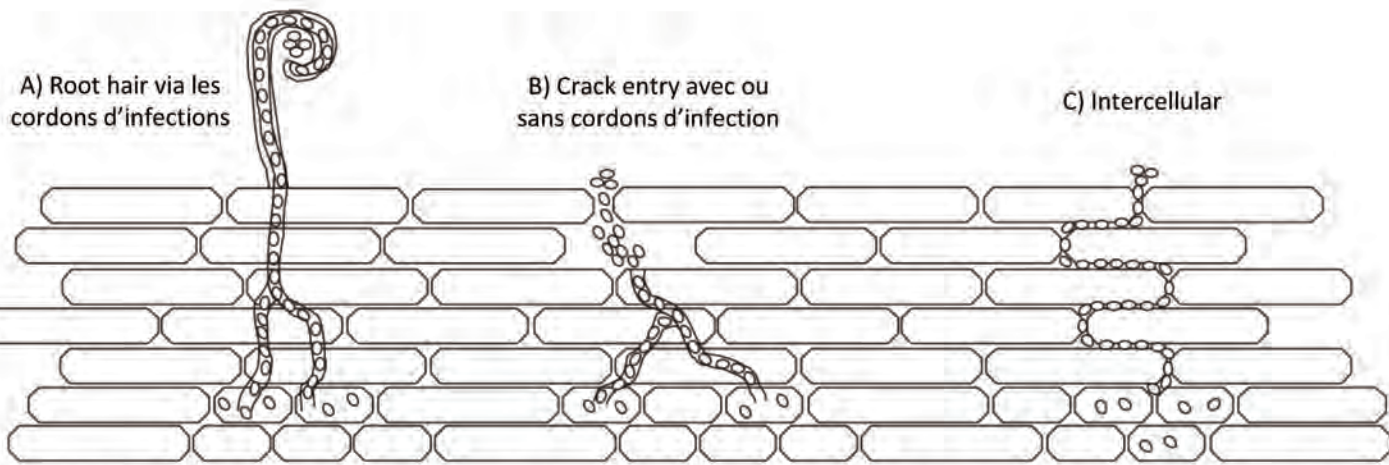


Figure I.2.1.4.1: Schéma représentant les différents modes d'infection des légumineuses par les rhizobia. Adapté de (Madsen et al. 2010).

les plantes modèles *M. sativa/truncatula* (Fournier et al. 2008) et *L. japonicus* (Madsen et al. 2010). Deux autres modes d'infection intercellulaires ont été identifiés, celui du « crack-entry » que l'on retrouve chez *Arachis hypogaea* (Chandler 1978) ou *Aeschynomene* (Bonaldi et al. 2011) et enfin le mode d'infection « intercellular » identifié chez *Mimosa scabrella* (De Faria et al. 1988) ou *Lupinus albus* (Gonzalez-Sama et al. 2004) (**Figure I.2.1.4.1**). Ces deux modes d'infections se différencient principalement par leur mode d'entrée dans la racine et dans la libération des bactéries dans le cortex.

Dans le cas de l'infection via « root hair », après attachement et multiplication des bactéries sur la racine, c'est au niveau du poil absorbant que l'infection va se produire. Le poil se recourbe sur lui-même à plus de 360° formant une structure dite en « crosse de berger » créant un site d'infection (Oldroyd and Downie 2008). La prolifération des bactéries au niveau du site d'infection conduit à un espace confiné où les bactéries sont enfermées dans ce qu'on appelle une chambre d'infection (Fournier et al. 2015). Les bactéries prolifèrent dans la chambre d'infection jusqu'à entraîner une lésion de la paroi végétale par hydrolyse. Ensuite, une invagination de la membrane plasmique va se former au sein du poil absorbant permettant aux bactéries d'avancer dans le poil absorbant (Gage 2004). Au fur et à mesure de l'avancée de l'invagination, et donc des bactéries, la plante dépose du matériel pariétal autour de l'invagination, appelé cordon d'infection (Gage 2002, 2004). Le cordon d'infection continue de progresser dans la cellule du poil où les bactéries continuent de se multiplier. Des dégradations localisées des parois végétales ainsi que l'apparition de ponts cytoplasmiques permettent la progression du cordon vers la racine. En parallèle, les cellules du cortex racinaire se divisent pour former ce qui est appelé, le primordium nodulaire. Une fois le primordium nodulaire atteint, le cordon d'infection se ramifie et pénètre les cellules de ce dernier. Les rhizobia sont alors relâchées, par un processus d'endocytose dans les cellules du primordium nodulaire, enrobées d'une membrane d'origine végétale qui les isole du cytoplasme (Gage 2004). Les bactéries endocytées forment alors le symbiosome, où elles arrêtent de se diviser pour se différencier en bactéroïdes et fixer l'azote. C'est au niveau de la paroi du symbiosome, où de nouveaux transporteurs sont exprimés et que les échanges entre la plante et les rhizobia vont avoir lieu dans le nodule nouvellement formé (**Figure I.2.1.4.2**).

Dans le cas de l'infection par « crack entry », les bactéries colonisent massivement la surface de la racine et pénètrent dans cette dernière via les espaces intercellulaires entre les poils racinaires et les cellules de l'épiderme dont la paroi est altérée (Chandler 1978; Bonaldi et al. 2011). Le processus d'infection cellulaire se fait alors par invagination de la paroi végétale similaire à une endocytose (Bonaldi et al. 2011). Après l'internalisation des bactéries, les cellules végétales infectées montrent des signes de mort cellulaire semblables à celles observées lors de réponses hypersensibles (HR) en cas d'infection par un pathogène. N'ayant plus d'intégrité structurelle, ces cellules collapseront et sont comprimées par la pression de turgescence des cellules corticales voisines et forment de diverticules qui permettent le transport des bactéries plus profondément dans le cortex (Bonaldi et al. 2011). Ensuite, les bactéries infectent les cellules corticales en contact direct avec les cellules collapsées.

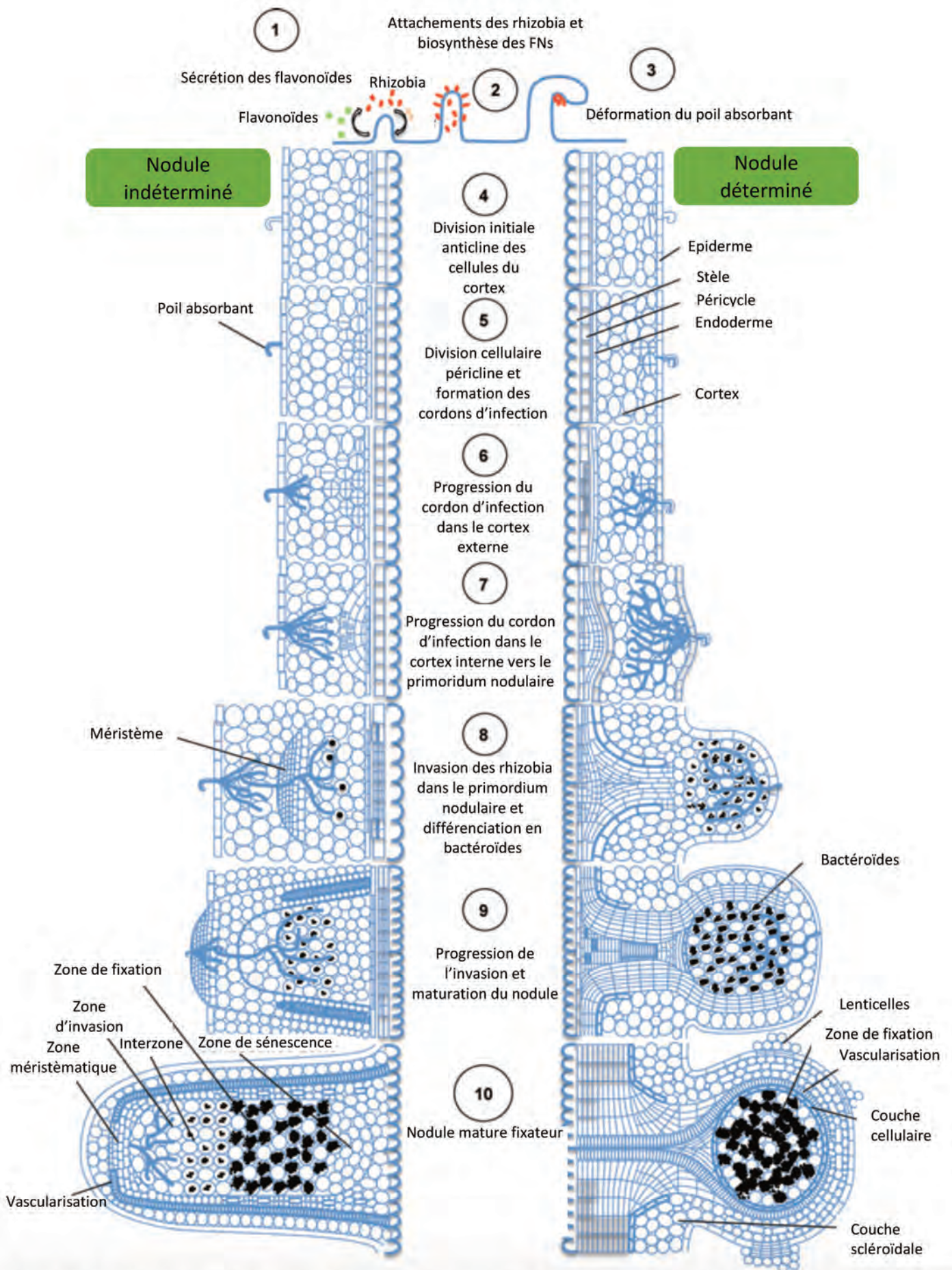


Figure I.2.1.4.2: Schéma chronologique de la mise en place d'un nodule lors d'une infection via root hair. Les flavonoïdes sont relâchés par la plante et reconnus par les rhizobia de la rhizosphère, qui vont alors synthétiser des facteurs Nod reconnus par la plante et permettre l'organogenèse du nodule. Adapté de (Oldroyd, 2013).

Néanmoins, les nouvelles cellules infectées conservent leur intégrité et se divisent pour donner naissance au nodule.

Dans le mode d'infection « intercellulaire », les bactéries pénètrent dans la racine végétale directement entre des cellules qui n'ont pas été altérées via des structures similaires aux cordons d'infections (De Faria et al. 1988; Gonzalez-Sama et al. 2004). Ici, les bactéries colonisent préférentiellement la racine principale au niveau des jonctions entre les poils absorbants et les cellules de l'épiderme adjacent. C'est entre ces deux cellules que les bactéries pénètrent dans la racine où elles vont se multiplier et infecter la cellule située sous le poil absorbant. L'invasion rhizobienne est concomitante avec l'induction des divisions cellulaires du cortex externe de la racine. L'accumulation des bactéries se fait ensuite dans la paroi des cellules du cortex située sous le poil absorbant. Ensuite, les bactéries sont libérées dans le cytoplasme des cellules du cortex via une altération de la paroi cellulaire qui permet l'endocytose des bactéries. Ensuite, la cellule infectée se divise rapidement et chacune des cellules fraîchement divisées contient un petit nombre de bactéries (Gonzalez-Sama et al. 2004). Après, la division des cellules du cortex externe s'accompagne de la division des cellules du cortex racinaire des couches plus profondes sous le site d'infection déterminant le primordium nodulaire. Le développement du primordium en nodule résulte de la division des cellules infectées et non infectées. Les rhizobia intracellulaires se divisent activement à l'intérieur des cellules hôtes. Dans ces cellules hôtes, les bactéries se situent à proximité de la paroi et sont réparties équitablement entre les cellules filles durant les étapes de division (Gonzalez-Sama et al. 2004).

Si généralement, les nodules sont retrouvés sur les racines, leur localisation peut varier et il existe quelques exemples de légumineuses capables de former des nodules caulinaires sur des racines adventives présentes sur la tige. C'est le cas par exemple de *Sesbania rostrata* et de certaines *Aeschynomene* (Masson-Boivin et al. 2009; Giraud and Fleischman 2004; Capoen et al. 2010). Deux types de structures nodulaires sont connues : indéterminés issus de la division de cellules corticales internes et déterminés issus du cortex externe (Vasse et al. 1990; Ndoye et al. 1994). Dans le cas des nodosités de type indéterminé on retrouve un méristème apical persistant qui entraîne la multiplication continue des cellules végétales à l'extrémité distale du nodule, ce qui lui donne une forme allongée. Au cours du développement nodulaire, les cellules végétales ne seront donc pas au même stade de développement au même moment résultant en différentes zones du nodule (Jones et al. 2007).

La région centrale du nodule indéterminé est différenciée en quatre ou cinq zones, selon l'âge du nodule (**Figure I.2.1.4.3**):

- La zone méristématique (zone I) est composée de petites cellules riches en cytoplasme.
- La zone d'infection (zone II) est constituée d'une quinzaine d'assises cellulaires (Vasse et al. 1990). C'est dans cette zone que les bactéries sont relâchées depuis les cordons d'infection, et sont internalisées dans le cytoplasme des cellules

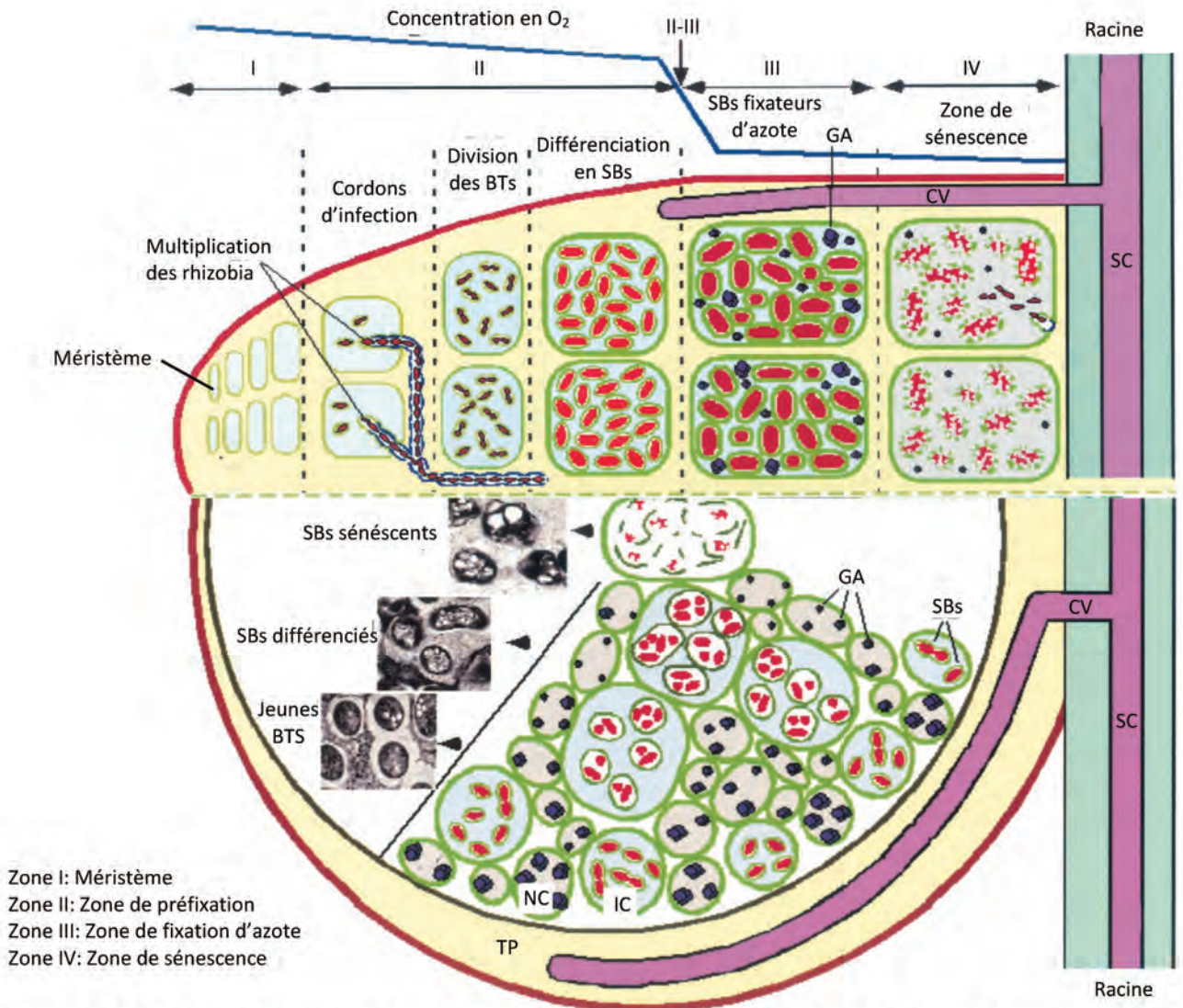


Figure I.2.1.4.3: Schéma représentant la morphogénèse des nodules de type indéterminés et déterminés. Comparaison de la morphogénèse d'un nodule indéterminés (en haut) et d'un nodule déterminés (en bas). SC: stèle centrale; CV: cordon vasculaire; BTs: bactéroïdes; SBs: symbiosomes; TP: tissu périphérique; IC: cellules infectées; NC: cellules non infectées; GA: grains d'amidons. Adapté de (Patriarca *et al*, 2002)

végétales par un processus d'endocytose. Dans la partie distale, les cellules végétales envahies subissent une différenciation morphologique, passant par des cycles d'endoréplication et un élargissement des cellules, pour devenir quasi sphériques dans la partie proximale (Timmers et al. 1998).

- L'interzone II-III s'étend sur une à trois assises cellulaires (Vasse et al. 1990). Dans cette zone, les cellules envahies sont remplies d'amyloplast. Avec le cortex nodulaire, cette zone constitue une barrière étanche qui empêche la diffusion d'oxygène vers la zone de fixation.
- La zone de fixation (zone III) est composée de cellules géantes polyploïdes en phase terminale de différenciation, et qui possèdent une grande vacuole centrale. Chez *Medicago*, ces cellules sont remplies de bactéroïdes disposés perpendiculairement à la membrane des cellules végétales (disposition en rayons de bicyclette) (Timmers et al. 1998). C'est dans cette zone qu'a lieu la conversion du diazote (N₂) en ammonium (Timmers et al. 1998).
- La zone de sénescence (zone IV) comprend les cellules de l'hôte et les cellules bactériennes en fin de cycle qui dégénèrent (Timmers et al. 1998).
- La zone saprophytique (zone V) contient des bactéries non-différenciées libérées des cordons d'infection restants, qui envahissent les cellules sénescents du nodule. Cette zone constituerait une niche écologique où les bactéries vivraient comme des organismes saprophytes.

Dans le cas des nodosités de type déterminé, le développement se fait à partir d'un méristème central, elles ont donc une forme sphérique. La croissance du nodule résulte d'une augmentation de la taille des cellules. Le cortex nodulaire est un tissu homogène composé de cellules végétales saines et de cellules infectées qui sont toutes au même stade de développement. La fixation se fait dans toutes les cellules infectées de la nodosité sans localisation particulière à un stade donné du développement (**Figure I.2.1.4.3**) (Patriarca et al. 2002).

I.2.2 Déterminants génétiques de la symbiose rhizobium-légumineuse

Les modes d'infections précédemment décrits peuvent être induit par différentes stratégies d'infection. La stratégie d'infection la plus connue et la plus étudié est celle impliquant les facteurs Nod mais de récentes études ont déterminés deux autres stratégies d'infection, celle utilisant le système de sécrétion de type 3 (SST3) et un mécanisme qui ne requiert pas le SST3 et *nod* indépendant qui n'est pas encore été élucidé (Giraud et al. 2007). Il est à noter qu'il n'existe aucun lien entre les différents modes d'infections et les stratégies de nodulation.

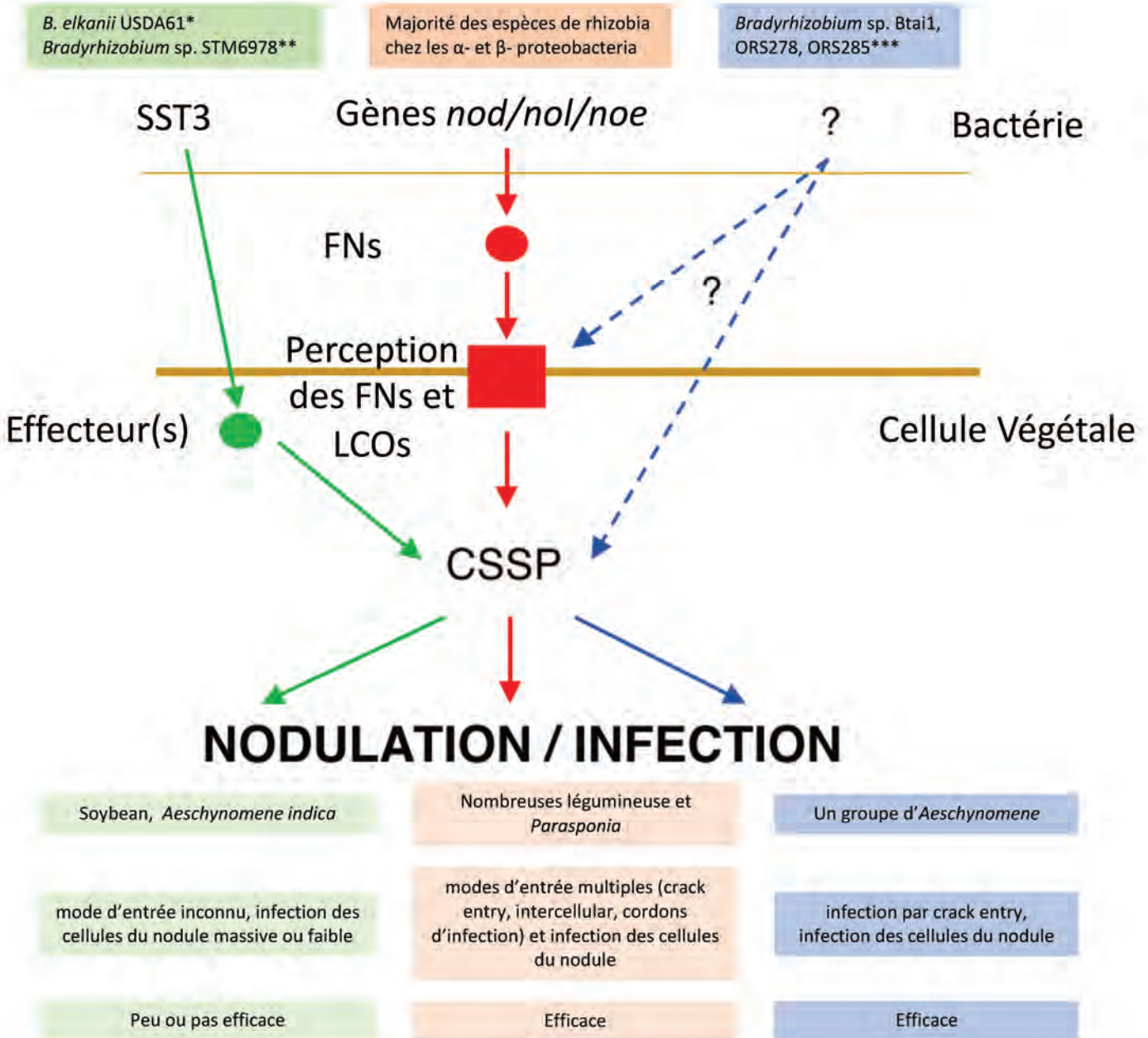


Figure I.2.2.1.1: Schéma récapitulatif des stratégies d'infection. *, Possède les stratégies Nod et SST3. ** Il n'a pas été démontré si l'utilisation du SST3 est l'unique stratégie de nodulation. ***, Possède les stratégies Nod, non-Nod et non-SST3. Adapté de (Masson-Boisin & Sachs, 2018).

1.2.2.1 Les stratégies de nodulation et d'infection

Pendant longtemps, la stratégie Nod a été considérée essentielle à la mise en place de la symbiose entre un rhizobium et une légumineuse et comme la clef de la spécificité symbiotique. Cette stratégie repose sur la production de lipochitooligosaccharides (appelés facteur Nods (FNs)) par les rhizobia et qui sont spécifiques à chaque souche. Ces FNs sont ensuite perçus par des récepteurs de plantes (LysM-receptor like kinases) qui activent la voie de signalisation symbiotique commune (common symbiosis signaling pathway ou CSSP). Les mécanismes génétiques et moléculaires impliqués dans la stratégie *nod* seront détaillés plus spécifiquement dans les parties 1.2.2.2 et 1.2.2.3.

La stratégie d'infection impliquant le SST3 a été mise en avant en laboratoire dans le système *Bradyrhizobium elkanii* USDA61 et *Glycine max* (Okazaki et al. 2013). Dans leur système, un mutant de USDA61 avec un SST3 défectueux qui produit des FNs est incapable d'induire des nodules (Okazaki et al. 2013). De plus, Okazaki et ses collaborateurs ont montrés qu'un mutant de USDA61 ne produisant pas de FNs mais qui possède un SST3 fonctionnel induit la formation de nodules non fonctionnels (Okazaki et al. 2013). Les effecteurs de type 3 (T3E) de nature et fonction symbiotique inconnues, mais en cours d'identification, permettent d'activer la CSSP en outrepassant les voies de reconnaissance des FNs (Okazaki et al. 2013, 2016; Songwattana et al. 2017). Le fait que les SST3 soient largement représentés chez les rhizobia, généralement sous le contrôle du régulateur NodD, et que le SST3 seul résulte en la formation de nodules non fonctionnels suggère que les stratégies *nod* et SST3 dépendante ont coopté au cours de l'évolution des rhizobia (Masson-Boivin and Sachs 2018).

Enfin, l'étude entre des légumineuses tropicales du genre *Aeschynomene* et des *Bradyrhizobium* photosynthétiques, a montré que cette interaction symbiotique n'implique pas la synthèse de facteurs *nod* par la bactérie. En effet le séquençage de deux souches de *Bradyrhizobium*, ORS278 et BTai1, en 2007, a révélé l'absence de gènes codant pour le squelette de base des FNs (*nodABC*) dans leurs génomes (Giraud et al. 2007) (**Figure 1.2.2.1.1**).

1.2.2.2 Le dialogue moléculaire de la symbiose FN dépendante

La symbiose qui s'établit entre la plante hôte et son symbiote rhizobien est le résultat d'interactions complexes entre les deux organismes que l'on qualifie de dialogue moléculaire. Le dialogue moléculaire commence lorsque la plante exsude, via ses racines, des composés phénoliques, les flavonoïdes, mais aussi des bêtaïnes ou encore des acides aldoniques (Perret et al. 2000). Les exsudats racinaires sont reconnus par des protéines régulatrices bactériennes à motif lysR (majoritairement NodD) qui déclenchent l'activation des gènes de nodulation bactérien (*nod*, *noe* et *nol*) (Spaink 2000; Perret et al. 2000). Cette reconnaissance représente un premier niveau de spécificité de la symbiose, les flavonoïdes étant spécifiquement reconnus par la protéine NodD.

L'activation des gènes de nodulation conduit à la biosynthèse des FNs, véritables molécules-signal, qui induisent le programme de développement végétales qui résulte en la formation

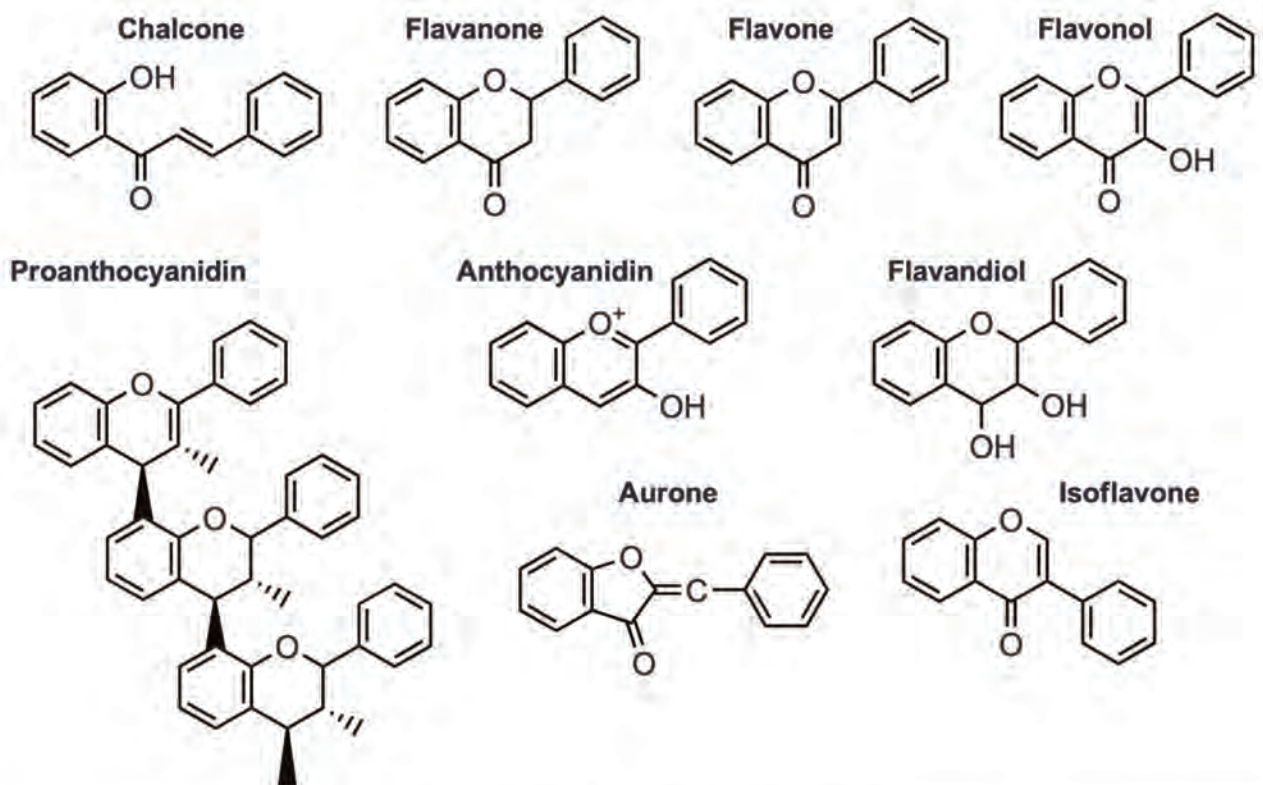


Figure I.2.2.1.1: Structures des classes majeures de flavonoïdes. Les flavonoides partageant tous une même structure de base formée par deux cycles aromatiques reliés par trois carbones : C₆-C₃-C₆, chaîne souvent fermée en un hétérocycle oxygéné hexa- ou pentagonal. présentent un squelette de base à 15 atomes de carbone, fait de deux cycles benzéniques C₆ reliés par une chaîne en C₃. Le pont à 3 carbones entre les deux phényles forme généralement un troisième cycle pyrone. (Falcone Ferreyra *et al*, 2012).

des nodules. La reconnaissance des FNs par la plante représente un deuxième niveau de spécificité. Ces deux niveaux de spécificité ne permettent pas à eux seuls d'expliquer la spécificité d'hôte qui existe entre les deux organismes symbiotiques.

I.2.2.2.1 Les flavonoïdes

A ce jour, plus de 5000 flavonoïdes naturels ont été identifiés chez les plantes vasculaires, néanmoins, seul un groupe particulier de flavonoïdes est connu pour être impliqué dans la symbiose fixatrice d'azote (Liu and Murray 2016; Ververidis et al. 2007). Ces composés phénoliques sont identifiés comme des métabolites secondaires de la famille des polyphénols et sont distribués en 7 sous-classes en fonction de leur structure et de leur degré d'oxydation : les aurones, les chalcones, les flavones, les flavonols, les flavandiols, les anthocyanines et enfin les proanthocyanidines (**Figure I.2.2.2.1.1**) (Falcone Ferreyra et al. 2012). Lors de la symbiose rhizobium-légumineuse, les flavonoïdes jouent des rôles multiples, notamment celui de chemoattractant sur les bactéries de la rhizosphère, puis ils vont permettre l'induction des gènes *nod* des rhizobia après avoir été reconnu par les protéines NodD compatibles. Enfin, ils interviennent aussi dans la régulation du développement nodulaire (Subramanian et al. 2007).

I.2.2.2.2 Reconnaissance des flavonoïdes et induction des gènes *nod*

Suite à la reconnaissance des flavonoïdes par NodD, le complexe flavonoïde/NodD provoque un changement de conformation de NodD qui peut alors se fixer sur des motifs de l'ADN bactérien de 47bp, appelées *nod*-boxes, situés en amont des clusters de gènes *nod* et activer leur transcription (Schlaman et al. 1989). Le nombre de copies de *nodD* ainsi que l'affinité qui existe entre NodD et les flavonoïdes est variable d'une souche à l'autre. Par exemple, *Sinorhizobium meliloti* possède 3 copies de *nodD* dont les produits reconnaissent peu de flavonoïdes différents (Broughton and Perret 1999). Au contraire, la souche *Sinorhizobium* sp. NRG234, qui possède un large spectre d'hôte, ne possède que 2 copies de *nodD* capables d'interagir avec de nombreux composés phénoliques (Perret et al. 1999). Le nombre de copies de *nodD* le plus élevé a été trouvé chez *Rhizobium tropici* CIAT899, qui sont au nombre de 5 (van Rhijn et al. 1993). Il s'avère que certaines copies de *nodD* codent pour des répresseurs de transcription des gènes *nod*. C'est le cas notamment de *nodD2* chez NGR234 (Fellay et al. 1998). D'autres régulateurs de transcription ont été identifiés, comme le gène *syrM*, un homologue de *nodD* que l'on retrouve chez *Sinorhizobium meliloti* et *Sinorhizobium* sp. NGR234. La protéine SyrM est impliquée dans la régulation *nod* mais permet aussi l'induction de la synthèse des exopolysaccharides (Maillet et al. 1990). Enfin, l'induction de l'expression des gènes *nod* chez *Bradyrhizobium japonicum* se fait par un système régulateur à deux composants, NodV-NodW (Sanjuan et al. 1994; Loh et al. 1997). La répression quant à elle se fait par des régulateurs transcriptionnels comme Nola qui permettent de réguler la mise en place de la symbiose en fonction des facteurs environnementaux (Gillette and Elkan 1996) ou par quorum sensing avec NwsA-NwsB (Loh et al. 2002). Certains facteurs environnementaux comme le pH ou la concentration en calcium peuvent influencer l'expression des gènes *nod*.

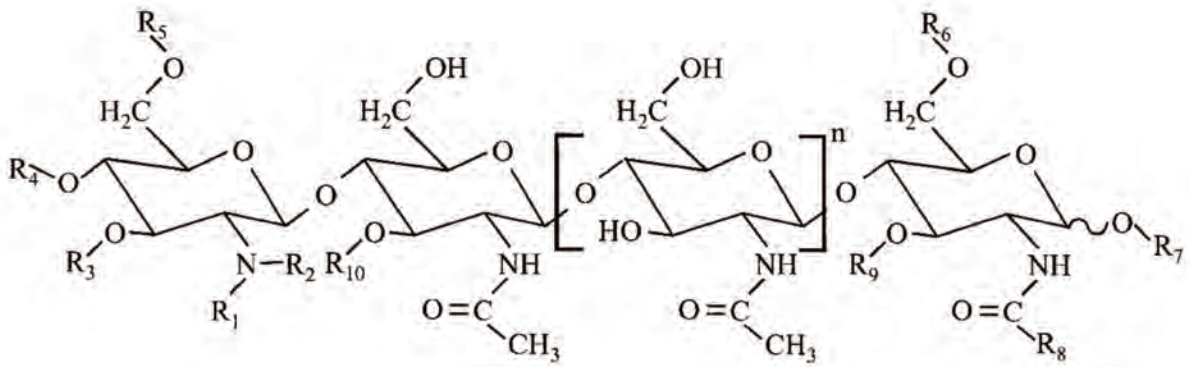


Figure I.2.2.3.1: Structure chimique des facteurs Nod majoritairement produits par les différentes espèces de rhizobia. Les FNs sont généralement formés d'une suite de 4 ou 5 unités de N-acétyl-D-glucosamine réunies par des liaisons $\beta(1 \rightarrow 4)$, et comprenant une chaîne acétel.

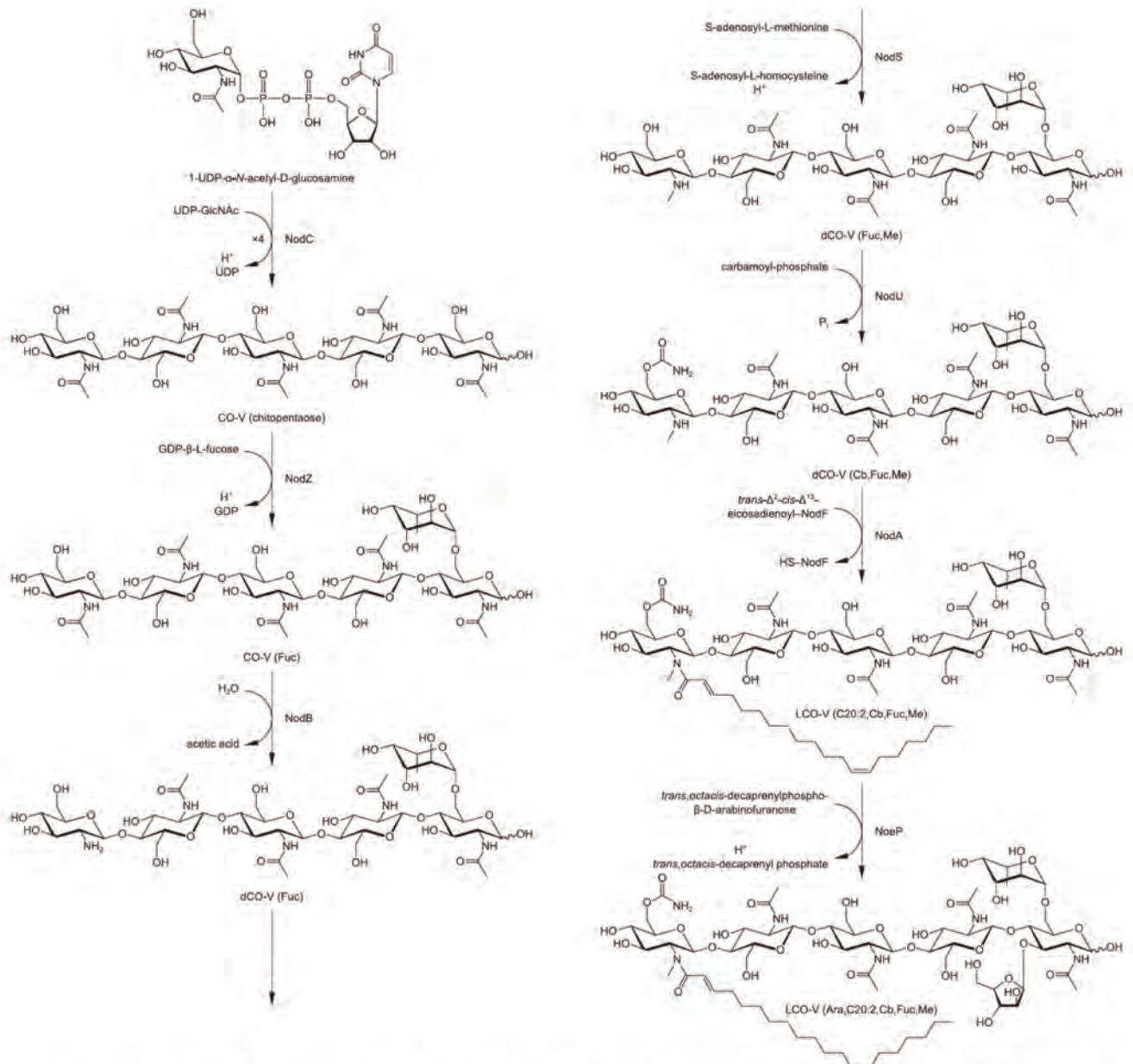


Figure I.2.2.3.2: Chaîne de biosynthèse des facteurs Nod, comme suggéré par les données MS. Exemple de la biosynthèse du FN majoritaire produit par *Rhizobium* sp. IRBG74 (Poinsot et al. 2016).

I.2.2.2.3 Les facteurs Nod

La plupart des FNs décrits actuellement présente un squelette composé de 3 à 6 résidus de N-acétyl-D-glucosamine (**Figure I.2.2.2.3.1**), substituée par une chaîne d'acyl au niveau de la glucosamine terminale et non réductrice (D'Haeze and Holsters 2002). Ce squelette est assemblé par les protéines NodABC. Les N-acétyl-D-glucosamine (GlcNaC) sont liés entre eux en β -1-4 par NodC qui est une N-actétylglucosamyl-transférase. Ensuite, la N-deacétylase NodB clive le résidu N-actétyl de la glucosamine terminale qui permettra ensuite d'y greffer une chaîne d'acide de 16 à 20 carbones. C'est NodA, une acyltransférase, qui permet le transfert de la chaîne d'acyl qui peut être commune (acides gras mono ou polyinsaturée provenant du métabolisme général des lipides), ou bien de type $\alpha\beta$ -mono/polyinsaturés (sous le contrôle de NodFE).

Selon le rhizobium, le squelette de base est décoré par des groupements chimiques variés, généralement présents sur les glucosamines situées aux deux extrémités grâce à la présence de set de gènes *nod* spécifiques (**Figure I.2.2.2.3.2**) (Perret et al. 2000; D'Haeze and Holsters 2002). Les modifications les plus fréquentes rencontrées sur les FNs sont des carbamoylations (*nodU, nolO*), glycosylations (*noeK, nolM, noeE*) sulfatations (*nodH, noeE*), acétylations (*nodL, nodX, nolL*), arabinosylations (*noeC*), méthylations (*nodS, noel*), fucosylations (*nodZ, nolK*) sur les glucosamines terminales avant d'être sécrété par NodI et NodJ, puis reconnus par la plante hôte adaptée (Perret et al. 2000). (Poinsot et al. 2016).

Il existe des FNs atypiques ou portant des substitutions rares. C'est le cas notamment de *Rhizobium gallicum* bv. *gallicum* R602 et *d'Ensifer* sp. BR816 qui synthétise des N-acétylglucosaminitol, c'est-à-dire des FNs dont le sucre réducteur a été réduit et forme un cycle ouvert (Snoeck et al. 2001; Soria-Díaz et al. 2006). Certains FNs de *Bradyrhizobium elkanii* portent un glycérol en position R7 qui représenterait un intermédiaire dans la chaîne de biosynthèse (Carlson et al. 1994). D'autres substitutions rares ont été retrouvés, notamment chez CIAT899 où un mannose est retrouvé en position R7 ou encore chez USDA191 dont certains FNs possèdent un glucose additionnel dans la chaîne glucosamine, alors que *M. loti* NZP2213 produit des composés qui portent un fucose en position R5.

I.2.2.3 Autres facteurs d'infection

D'autres facteurs sont requis pour le bon déroulement des étapes précoces et tardives du processus d'infection depuis le développement du cordon d'infection jusqu'à l'entrée des rhizobia dans les cellules du nodule.

I.2.2.3.1 Les polysaccharides de surface

Chez les rhizobia, quatre types de polysaccharides de surface principaux ont pu être identifiés : les exopolysaccharides (EPS), les lipopolysaccharides (LPS), les polysaccharides capsulaires (KPS) et enfin les β -glucanes cycliques (GC) (Cooper 2007). Ces différents polysaccharides ont été démontrés pour être impliqués dans la colonisation de la rhizosphère,

la reconnaissance du symbiote par l'hôte mais aussi dans la persistance des cordons d'infection et l'invasion de la nodosité. Même si la synthèse des polysaccharides n'est pas dépendante des flavonoïdes, ces derniers peuvent influencer leur structure (Cooper 2007).

Les EPS sont des polysaccharides homo- ou hétérométrique qui sont sécrétés à la surfaces des bactéries. La composition des EPS dépend de leur degré de polymérisation ainsi que de la nature des monomère qui les composent, D-glucose, D-galactose ou bien D-mannose (Skorupska et al. 2006). Les EPS peuvent aussi contenir d'autres composés tel que de l'acétate, du pyruvate ou encore du succinate (Skorupska et al. 2006). Le rôle des EPS lors de la symbiose fixatrice d'azote rhizobium-légumineuse a été principalement étudié chez *Sinorhizobium meliloti* et *Rhizobium leguminosarum* qui nodulent la luzerne et le pois, respectivement. Il a été démontré que des mutants affectés pour la synthèse des EPS chez ces deux rhizobia auraient une capacité d'infection altérée (Breedveld et al. 1993; Cheng and Walker 1998).

Les LPS sont des composants majeurs de la membrane externe des bactéries Gram négatives (Alexander and Rietschel 2001). Ces molécules se composent de trois régions structurales distinctes: une région composée de polysaccharides, dont la structure est très variable, appelée antigène-O qui est dirigée vers l'extérieur de la cellule, une région composée d'oligosaccharides et appelée core oligosaccharidique permettant le lien entre la région antigène-O et la partie lipidique des LPS appelée lipide A qui, elle-même, permet l'encrage de cette molécule dans la membrane externe des bactéries (Alexander and Rietschel 2001). Les LPS produits par les rhizobia permettent la suppression des réponses immunitaires de la plante (Jones et al. 2007). Certains mutants, dont la structure des LPS est altérée, présentent un retard pour la nodulation (Brown et al. 2011), une senescence précoce des nodules ou encore un processus de différenciation en bactéroïdes altérées (Haag et al. 2009).

On retrouve les KPS, aussi appelés antigène-K, chez de nombreux rhizobia. Ces polysaccharides capsulaires qui peuvent varier d'une souche à une autre au sein d'une même espèce (Forsberg and Reuhs 1997), seraient impliqués dans les étapes précoces de l'infection (Becquart-de Kozak et al. 1997). En effet, il semblerait que l'ajout exogène de KPS de *S. meliloti* provoque l'exsudation de flavonoïdes chez la luzerne (Becquart-de Kozak et al. 1997). Les voies de biosynthèse des KPL, EPS et LPS étant partiellement communes, il difficile de déterminer le rôle précis des KPL dans la symbiose.

Les GC sont des homopolymères de glucose de faible poids moléculaire que l'on retrouve chez les rhizobia dans leur espace périplasmique ou sécrétés dans le milieu extracellulaire (Breedveld and Miller 1994). Différentes études faites sur des mutants affectés dans la synthèse des GC ont montrés leur incapacité à infecter leur hôte. Ces résultats suggèrent que les GC sont impliqués dans les étapes précoces de la nodulation (Bhagwat et al. n.d.; Crespo-Rivas et al. 2009; Breedveld and Miller 1994).

I.2.2.3.2 Les systèmes de sécrétion

Chez les rhizobia, plusieurs protéines de sécrétion sont connues pour jouer un rôle dans la symbiose. En effet, elles vont permettre de supprimer les réactions de défense de la plante ou influencer le spectre d'hôte (Cooper 2007; Vlassak et al. 1998). Ces protéines de sécrétion sont transportées via des systèmes de sécrétion d'utilisation générale comme le système de sécrétion de type 1 (T1SS) mais aussi par des systèmes de sécrétion plus spécialisés comme les systèmes de sécrétion de type III, IV, et VI (SST3, SST4 et SST6) (Cooper 2007; Fauvart and Michiels 2008).

Le T1SS est notamment connu pour sécréter la protéine NodO qui est encodée par le gène *nodO*, inductible par les flavonoïdes et sous contrôle de NodD. NodO est une protéine de type calcium-binding qui permet le développement des cordons d'infection dans les poils absorbants (Walker and Downie 2000). Le transfert de *nodO* chez d'autres rhizobia permet d'augmenter son spectre d'hôte (Vlassak et al. 1998).

Le rôle du SST3 dans la symbiose est celui qui a été le plus étudié. On le retrouve chez la plupart des rhizobia et est responsable de la sécrétion de Nops (nodulation outer proteins) (Krishnan et al. 2003). Contrairement au SST1, il est fréquent que l'expression des gènes codants pour le SST3 et les protéines sécrétées par ce dernier soit dépendante des flavonoïdes et de NodD (Cooper 2007). La plupart des clusters de gènes codant pour le SST3 ne possèdent pas de *nod*-box (Cooper 2007). L'expression de ces gènes est dépendante d'autres motifs de régulations dans leur région promotrice que l'on appelle les *tts*-box qui sont régulées par TtsI (Cooper 2007). TtsI est un facteur de transcription encodé par le gène *ttsI*, qui est lui-même sous contrôle d'une *nod*-box et régulé par les flavonoïdes et NodD (Cooper 2007). De façon générale le SST3 et les protéines Nops en particulier peuvent avoir une influence positive, négative ou nulle sur la nodulation en fonction de la plante hôte impliquée (Cooper 2007). On sait par exemple que l'inactivation du SST3 chez *C. taiwanensis* d'augmente son spectre d'hôte (Saad et al. 2012). De nombreux Nops putatifs ont été identifiés chez les rhizobia, notamment NopL, qui semblerait être impliqué dans la suppression des réactions de défenses de la plante comme c'est le cas pour le SST3 des bactéries pathogènes (Bartsev et al. 2004).

I.2.2.4 Voies de signalisation végétales conduisant à l'infection et à l'organogénèse du nodule

L'étude des plantes modèles *Medicago truncatula* et *Lotus japonicus* et de leurs différents mutants ont permis d'identifier les acteurs moléculaires impliqués dans la voie de signalisation en réponse aux FNs rhizobiens. Ainsi, de nombreux gènes et protéines impliqués dans la perception des FNS et dans la transduction du signal en réponse à ces derniers ont pu être identifiés.

Lotus japonicus

Medicago truncatula

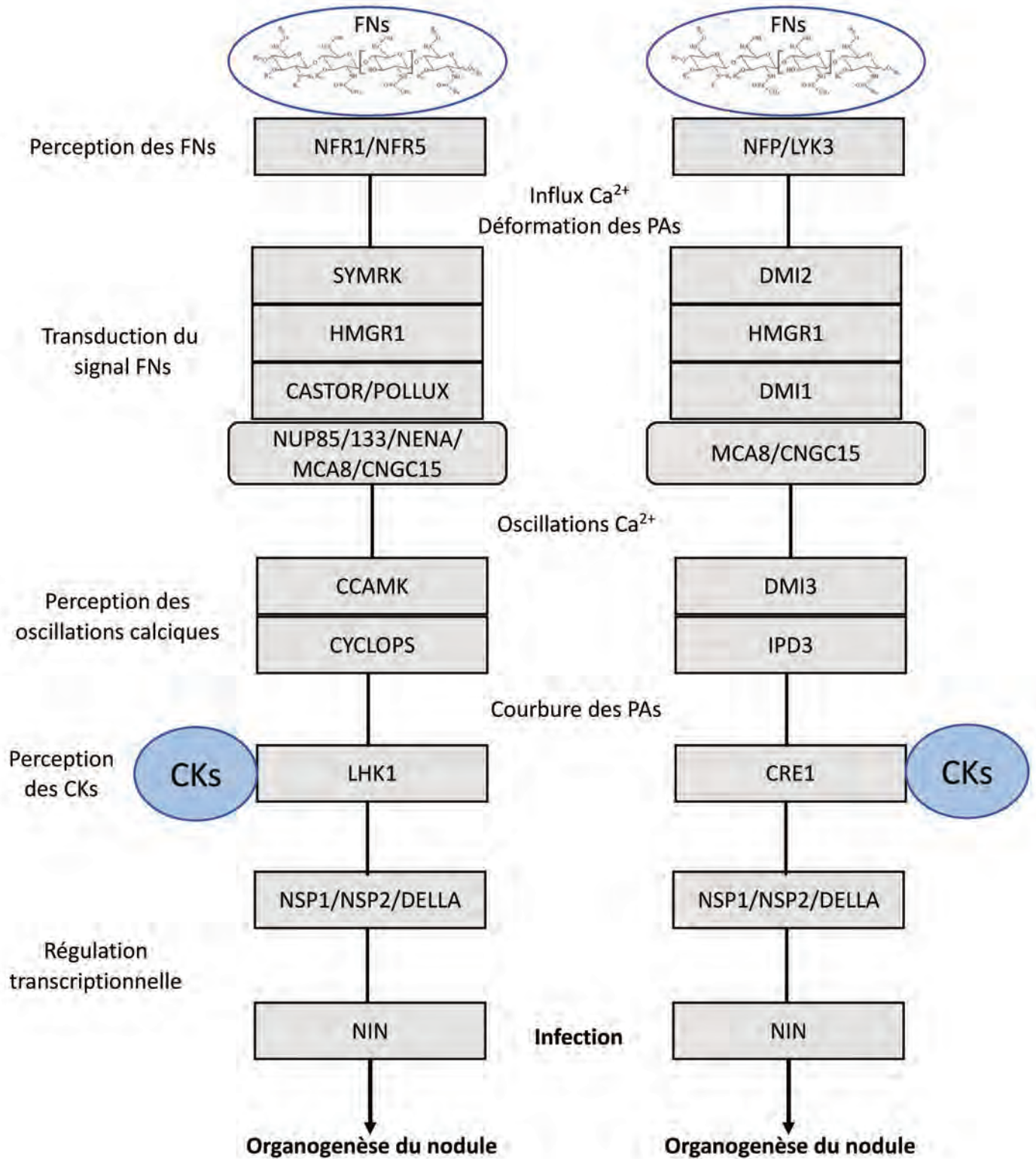


Figure 1.2.2.4.2.1: Modèle simplifié des acteurs moléculaires impliqués dans la voie de signalisation FNs chez *L. japonicus* et *M. truncatula*. Les protéines orthologues sont positionnées en face les unes des autres. FNs: facteurs Nods; PAs: poils absorbants; CKs: cytokinines

I.2.2.4.1 Perception des NFs

Des récepteurs sérine/thréonine kinase à domaine LysM extracellulaire, les LysM-RLKs, ont été identifiés comme récepteurs putatifs aux NFs chez *Medicago truncatula* (NFP/LYK3) et *Lotus japonicus* (NFR1/NFR5) (Amor et al. 2003; Madsen et al. 2003; Oldroyd and Long 2003). Les LysM-RLKs se présentent sous forme d'hétérodimères où NFR1 est couplé à NFR5 et LYK3 avec NFP. Les domaines kinases de NFR1 et LYK3 sont fonctionnels et permettent la transduction du signal des NFs. Les domaines kinase de NFR5 et NFP sont quant à eux incomplets et donc dépourvus d'activité de phosphorylation. Néanmoins, l'inactivation de NFR5 et LYK3 ne permet pas la nodulation ce qui indique que l'hétérodimérisation est indispensable pour fonctionner (Nakagawa et al. 2011; Madsen et al. 2011). Cette première étape induit des influx calciques au sein de la membrane plasmique et la déformation des poils absorbants. La voie de signalisation NFs peut être divisée en deux parties synchronisées. La première conduit à l'infection et se déroule au sein de l'épiderme et la seconde induit l'organogenèse du primordium nodulaire dans le cortex.

I.2.2.4.2 Transduction du signal NFs

Après que les NFs soient reconnus par les récepteurs LysM-RLKs plusieurs protéines interviennent dans la production d'oscillations calciques nucléiques. La **Figure I.2.2.4.2.1** représente ces différentes protéines ainsi que l'ordre de la voie de signalisation.

La première protéine de cette voie de signalisation, SYMRK chez *Lotus* et DMI2 chez *Medicago*, est une récepteur kinase à domaine Leucine-Rich-Repeat (LRR-RLK) qui se situe dans la membrane plasmique impliqué dans la transduction du signal des NFs en aval des LysM-RLKs (Endre et al. 2002; Miwa et al. 2006).

Les protéines CASTOR et POLLUX (DMI1 chez *M. truncatula*), MCA8 et CNGC15 sont directement impliquées dans la production des oscillations calciques. CASTOR et POLLUX/DMI1 sont des canaux ioniques localisés dans l'enveloppe nucléaire (Riely et al. 2007) qui permettent des transferts d'ions potassium (K^+) entre le cytosol et l'intérieur du noyau (Charpentier et al. 2016). Les flux de potassium permettent de conserver un équilibre des charges au sein du noyau lors des oscillations calciques induites par MCA8, une calcium-ATPase, et les protéines CNGC15 qui forment un canal calcique (Capoen et al. 2011; Charpentier et al. 2016).

Ces oscillations calciques ont un rôle de messenger secondaire perçu par DMI3 (*M. truncatula*) et CCaMK (*L. japonicus*), des protéines kinase calcium et calmoduline-dépendante qui se trouvent dans le noyau des poils absorbants (Capoen et al. 2011). DMI3/CCaMK est maintenu sous forme inactive par autophosphorylation de la protéine lorsque du calcium est directement fixé à cette dernière (Lévy et al. 2004; Takeda et al. 2012). Lors des oscillations calciques, le calcium se lie à la calmoduline qui entraîne l'inhibition de l'autophosphorylation de la protéine permettant la phosphorylation du facteur transcriptionnel CYCLOPS (*L. japonicus*) ou IPD3 chez *M. truncatula* (Messinese et al. 2007; Yano et al. 2008). CYCLOPS/IPD3

fait partie d'un complexe protéique qui comprends d'autres facteurs de transcription à domaine GRAS (NSP1, NSP2 et DELLA) qui, après phosphorylation par CCaMK ou DMI3, induit l'expression des gènes de nodulation précoce de la plante nécessaire à la formation et à la progression du cordon d'infection mais aussi à l'organogénèse nodulaire.

I.2.2.4.3 Libération des rhizobia et différenciation en bactéroïdes

La progression du cordon d'infection permet de conduire les bactéries au primordium nodulaire situé dans le cortex. Lorsque le cordon d'infection atteint le primordium nodulaire, les bactéries sont libérées dans les cellules du primordium par endocytose via une vésicule appelée « infection droplet » (Perret et al. 2000). Cette nouvelle structure, appelée symbiosome, comprend des bactéries qui se différencient en bactéroïdes la membrane végétale dans laquelle ils sont enfermés (Huisman et al. 2012). En fonction de l'hôte, les changements métaboliques qui s'opèrent chez la bactérie lors de sa différenciation en bactéroïde peuvent s'accompagner de changements en terme de taille et de forme (Haag et al. 2013). Par ailleurs, dans les nodules déterminés, les symbiosomes individuels fusionnent et/ou les bactéroïdes se divisent au sein même du symbiosome résultant en des symbiosomes qui peuvent contenir plus plusieurs bactéroïdes. Au contraire, dans les nodules de type indéterminés les symbiosomes contiennent généralement un unique bactéroïde (Prell and Poole 2006), à l'exception de *M. pudica* dont les symbiosomes contiennent jusqu'à 4 bactéroïdes (Marchetti et al. 2011).

Chez les plantes du genre *Aeschynomene* et celles appartenant au clade des Inverted Repeat-Lacking Clade (IRLC), comme *Medicago*, les bactéroïdes différenciés sont incapables de reprendre leur forme d'origine et de se diviser (Mergaert et al. 2006; Kereszt et al. 2011; Czernic et al. 2015). Cette différenciation est appelée « différenciation terminale des bactéroïdes » (ou TBD pour terminal bacteroid differentiation) et est contrôlée par la plante via des peptides antimicrobiens cationiques riches en cystéines nommées NCRs (Nodules cysteine-riche peptides) mais aussi par la présence de BacA chez les bactéries (Mergaert et al. 2003).

I.2.3 Évolution de la symbiose fixatrice d'azote

La plupart des plantes, incluant mousses, fougères, gymnospermes et angiospermes sont capables de former des associations lâches avec des bactéries fixatrices d'azote. Cependant seul un petit nombre d'angiospermes ont évolué la capacité d'héberger des fixateurs d'azote de manière intracellulaire, permettant un contrôle de leur prolifération et des échanges métaboliques. Parallèlement, seuls quelques genres bactériens ont évolué la capacité à fixer l'azote en symbiose. Comment ces symbiotes et ces interactions ont évolué est une question fascinante qui suscite un regain d'intérêt ces dernières années.

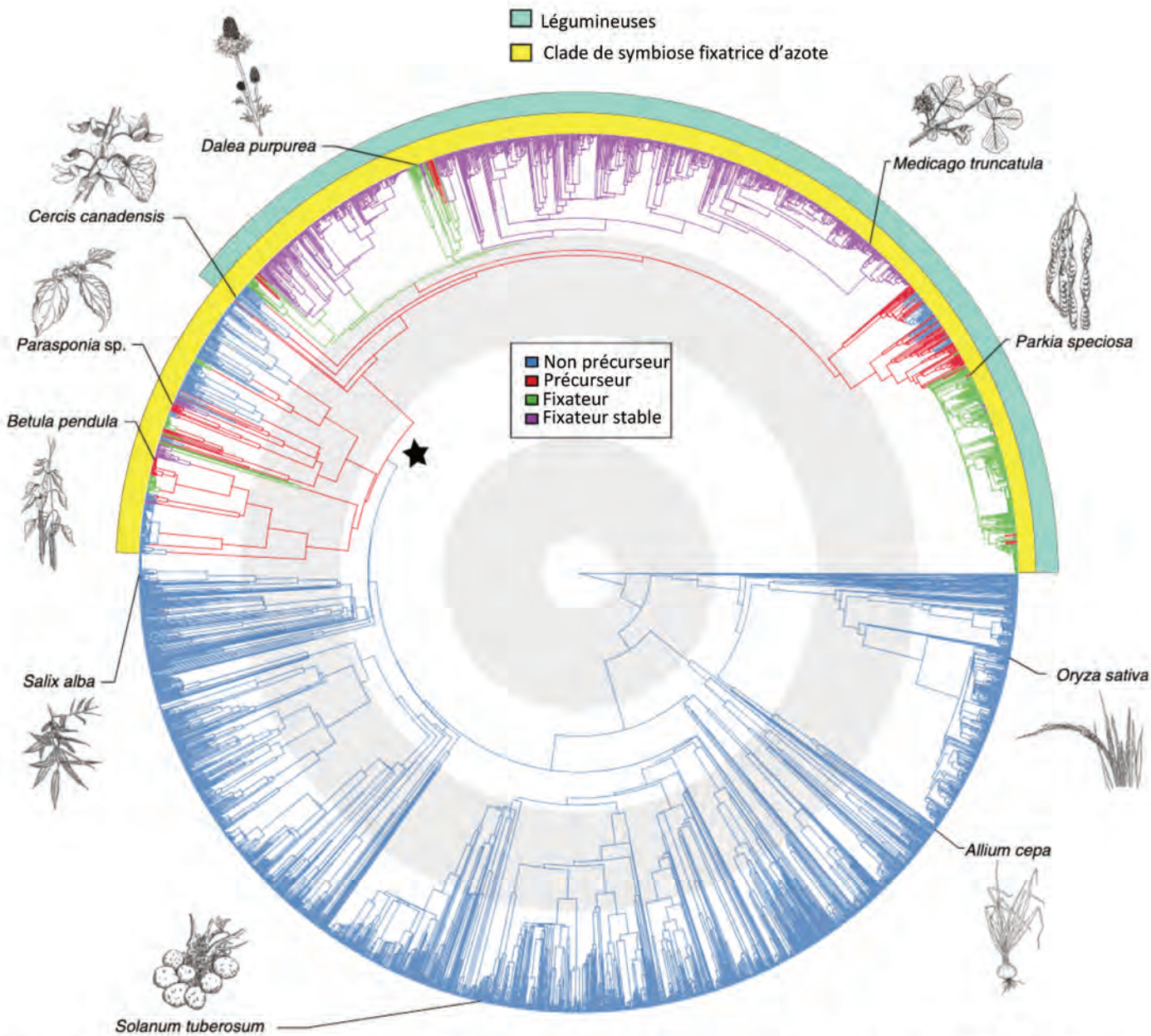


Figure 1.2.3.1.1: Phylogénie des angiospermes de 3467 espèces montrant la reconstruction de l'état nodulant. Les branches sont colorées en fonction de l'état le plus probable de leur ancêtre nodulant. L'étoile indique le précurseur d'origine. Le turquoise et le jaune indiquent les légumineuses du clade de la symbiose fixatrice d'azote rhizobium-légumineuse, qui contient tous les angiospermes nodulants connus. Les cercles gris et blancs indiquent les périodes de 50 millions d'années à partir du présent. Les dessins représentent des angiospermes importants. Adapté de (Werner et al. 2014).

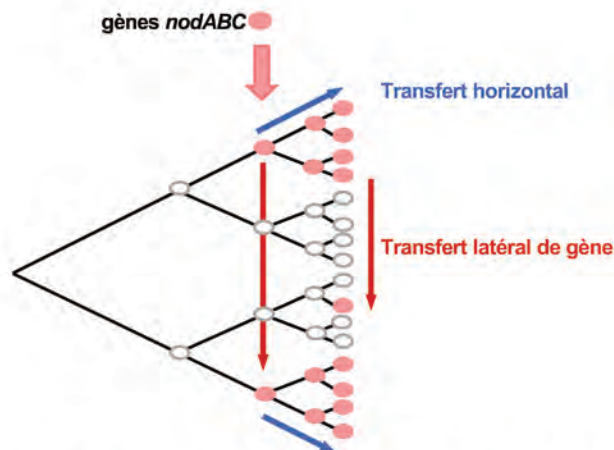


Figure 1.2.3.2.1: Schéma du scénario évolutif permettant d'expliquer l'origine polyphylétique des rhizobia. Les cluster de gènes *nod*, localisés sur des éléments génétiques mobiles, peuvent être transmis par transfert horizontal gènes à diverses bactéries du sol.

I.2.3.1 Les plantes

La réorganisation de l'arbre phylogénétique des angiospermes en 1995 a révélé que les plantes capables de former des nodules appartenaient aux ordres des Fabales, Fagales, Curbitales et Rosales (FaFaCuRo) (Soltis et al. 1995). Au sein des FaFaCuRo, seulement 10 des 28 familles contiennent des espèces capables de former des nodules. Ces familles ne forment pas un groupe monophylétique (Doyle 2011). Huit familles de l'ordre des Fagales, Rosales et Curcubitales contiennent des espèces nodulées par des *Frankia* (Pawlowski and Demchenko 2012). Les légumineuses (Fabales) et *Parasponia* (Rosales) sont capables de former des nodules avec des rhizobia (Akkermans et al. 1978). Chez les légumineuses, la capacité à noduler est variable : 90% des *Mimosoideae* et *Papilionoideae* sont nodulés contre 30% des *Caesalpinioideae*.

La distribution phylogénétique des plantes nodulées au sein des FaFaCuRo a conduit dans un premier temps à l'hypothèse que la symbiose nodulaire avait évolué en deux étapes : un premier événement de prédisposition chez un ancêtre commun, suivi par de multiples événements indépendants d'évolution de la nodulation (théorie de l'origine multiple) (**Figure I.2.3.1.1**) (Soltis et al. 1995; Swensen 1996; Doyle 2011; Werner et al. 2014).

Deux études récentes de génétique et génomique comparative d'espèces nodulées et non nodulées proposent aujourd'hui un autre schéma : la nodulation serait apparue il y a environ 100 millions d'années chez un ancêtre des FaFaCuRo et aurait été perdue dans de nombreuses lignées descendantes (van Velzen et al. 2018; Griesmann et al. 2018).

I.2.3.2 Les bactéries

A l'exception des *Bradyrhizobium* photosynthétiques, dont les mécanismes de nodulation sont inconnus à ce jour (Giraud et al. 2007; Tang and Capela 2020), tous les rhizobia utilisent la stratégie *nod* pour initier la nodulation. Cette stratégie quasi-universelle est bien connue même si certains rhizobia possèdent un SST3 qui peut partiellement se substituer aux gènes *nod* (Okazaki et al. 2013).

I.2.3.2.1 Origine par transfert horizontal de gènes

Les rhizobia dont la symbiose dépend de la stratégie *nod* auraient évolué à partir du transfert horizontal (HGT) d'un set de gènes symbiotiques (gènes *nod* et gènes *nif*) dans différents taxons (**Figure I.2.3.2.1**). Cette hypothèse repose sur les arguments suivants :

- 1) les gènes *nod* sont situés sur des éléments génétiques mobiles (plasmide, îlots symbiotiques) ce qui facilite leur transfert.
- 2) les rhizobia forment un groupe polyphylétique (Remigi et al. 2016) alors que les gènes *nod* canoniques (*nodABC*) ont une origine monophylétique chez les rhizobia (Persson et al. 2015).
- 3) Les différents genres de rhizobia existaient déjà avant l'apparition des légumineuses il y a 70-100 millions d'années (Turner and Young 2000).

Cette hypothèse a été confirmée par les travaux de Sullivan *et al.* qui ont montré en laboratoire et au champ que la transmission d'îlots symbiotiques par transfert horizontal était possible (Sullivan *et al.* 1995; Sullivan and Ronson 1998). Dans une expérience célèbre, la légumineuse *Lotus corniculatus* et son inoculum *Mesorhizobium loti* ICMP3153 ont été introduits dans un environnement naturel dépourvu de symbiotes de *Lotus*. Sept ans plus tard, une étude de la diversité génétique des symbiotes de *L. corniculatus* dans cet environnement a montré une grande diversité génétique des bactéries isolées des nodules de *L. corniculatus*. De façon intéressante, tous les isolats possédaient le même îlot symbiotique que ICMP3153 (Sullivan *et al.* 1995). Il a par la suite été démontré que le transfert de l'îlot symbiotique de ICMP3153 vers les autres bactéries du sol est à l'origine de cette diversité génétique (Sullivan and Ronson 1998).

A ce jour, on suppose que les gènes *nod* présents chez les rhizobia ont été à l'origine acquis de *Frankia* (Persson *et al.* 2015; Nguyen *et al.* 2016; Ktari *et al.* 2017). L'analyse du génome *Candidatus* *Frankia daticae* Dg1 (Dg1), a révélé l'existence des gènes canoniques *nodABC*. Ces gènes s'organisent en deux opérons différents et sont exprimés dans les nodules de la plante hôte, *Dastica glomerata* (Persson *et al.* 2015). Le gène *nodC* de Dg1 complémente partiellement un mutant de *Rhizobium leguminosarum* A34 *nodC*::Tn5 et l'analyse phylogénétique des protéines NodABC de Dg1 a montré que ces dernière se situent à la racine des protéines NodABC des α - et β -rhizobia (Persson *et al.* 2015). Les analyses de génome d'autres souches de *Frankia* ont permis de déterminer l'existence des gènes *nodABC* dans les cluster II et III (Persson *et al.* 2015; Nguyen *et al.* 2016; Ktari *et al.* 2017). De façon intéressante, chez les protéobactéries, on ne retrouve *nodA* que chez les rhizobia alors que ce gène est fréquent chez les actinobactéries, y compris les non symbiotiques comme *Streptomyces bottropensis* (Persson *et al.* 2015).

La spécialisation des gènes *nod* dans la symbiose pourrait être issue d'une néofonctionnalisation de gènes paralogues impliqués dans d'autres fonctions. Par exemple, il semblerait que *nodB* et *nodC* aient évolués à partir de gènes codant pour des protéines de fonctions similaires présents chez d'autres organismes. On retrouve des homologues de la N-deacétylase NodB chez des bactéries qui ne sont pas des fixateurs symbiotiques, comme *Streptomyces lividans* et *Clostridium thermocellum* (Taylor *et al.* 2006). De la même manière, NodC appartient à une classe importante de β -glucosyl transférases qui sont connues pour incorporer des résidus N-acétyl-glucosamine dans les parois bactériennes. D'autres protéines de cette famille sont impliquées dans la production de peptidoglycanes en assemblant les résidus d'acétyl-glucosamines et certains de ses dérivés, notamment les N-acétyl-acide muramique (Hirsch *et al.* 2001). Il n'est pas impossible que *nodC* soit le résultat d'un processus de divergence fonctionnelle d'une de ces enzymes.

L'origine des premiers rhizobia fait encore débat puisque pour certains, les premiers rhizobia appartiendraient au genre *Bradyrhizobium* (Martínez-Romero *et al.* 2010; Ormeño-Orrillo *et al.* 2013) pour d'autres, les premiers rhizobia viendraient plutôt des β -protéobactéries (Aoki *et al.* 2013).

I.2.3.2.2 Étape d'optimisation post-HGT

Le fait que la symbiose fixatrice d'azote soit restreinte à une quinzaine de genres bactériens est surprenant étant donné la grande diversité taxonomique des populations bactériennes de la rhizosphère et la présence de bactéries non symbiotiques au sein des nodules (Remigi et al. 2016). Dans la mesure où les transferts de gènes sont extrêmement fréquents dans la nature, ces éléments suggèrent que l'acquisition des gènes symbiotiques par HGT n'est pas suffisant pour transformer une bactérie du sol en symbiote de légumineuse (Remigi et al. 2016). Ceci a été démontré au champ par une étude faite sur la légumineuse *Biserrula pelecinus* L.. Cette légumineuse de la région méditerranéenne a été implantée en Australie où les bactéries indigènes n'étaient pas capables de noduler cette plante (Nandasena et al. 2006). Six ans après l'introduction de *Biserrula pelecinus* L. et de son inoculum *Mesorhizobium ciceri* bv. *biserrulae* WSM1271 en Australie, une étude a révélé la présence d'au moins sept isolats différents dans les 88 nodules étudiés issus du transfert horizontal de l'îlot symbiotique de *M. ciceri* bv. *biserrulae* WSM1271 (Nandasena et al. 2006). Cinq de ces nouveaux rhizobia était peu fixateurs et les deux autres ne fixaient pas du tout montrant que le mutualisme n'était pas achevé (Nandasena et al. 2007). La conséquence de ces événements de transferts multiples et indépendants est donc l'évolution de bactéries métaboliquement et génétiquement variées. La comparaison de 7 génomes d' α - et β -protéobactéries a montré l'existence d'un set de gènes communs mais dont aucun n'est commun et spécifique au rhizobia (Amadou et al. 2008). Il est donc difficile de savoir aujourd'hui si les genres qui ne contiennent pas de rhizobia sont capables de le devenir ou si il s'agit d'une question de chance ou de hasard (Amadou et al. 2008; Remigi et al. 2016).

Il a par la suite été démontré en laboratoire que le HGT n'est que la première étape pour qu'une simple bactérie du sol devienne un rhizobium fixateur d'azote (Marchetti et al. 2010). L'optimisation post HGT du potentiel symbiotique par le remodelage du génome d'accueil sous pression de sélection de la plante a été démontré par évolution expérimentale (Marchetti et al. 2010; Guan et al. 2013; Capela et al. 2017). Ces travaux sont détaillés dans la partie I.3.3.

Les données phylogénétiques, génétiques et expérimentales indiquent que l'apparition d'un nouveau genre de rhizobium est rare et que plus le donneur et le receveur sont phylogénétiquement éloignés plus l'étape d'optimisation est requise (Remigi et al. 2016). Il est probable que lorsque les gènes sont transférés chez une espèce proche, le receveur a des antécédents de symbiose dus à la perte/acquisition fréquente des traits symbiotiques (Remigi et al. 2016). C'est ce qu'on retrouve par exemple chez les espèces possédant plusieurs symbiovars (Rogel et al. 2011). Ceci a notamment été démontré avec la conversion de la souche de *R. leguminosarum* symbiovar *phaseoli* en un symbiovar *viciae* complètement fonctionnel simplement en remplaçant le plasmide symbiotique de *phaseoli* par celui du symbiovar *viciae*, sans réarrangement du génome d'accueil (Johnston et al. 1978).

1.2.3.3 Évolution du mutualisme

La fixation de l'azote atmosphérique par les bactéroïdes est l'étape ultime de la symbiose. Beaucoup de rhizobia ne réalisent pas cette étape avec certains hôtes compatibles pour la nodulation. Ils peuvent noduler et infecter sans être capable de fixer l'azote de l'air dans les cellules végétales. La capacité à noduler et infecter les nodules sans fixer l'azote ouvre la possibilité d'évoluer des « cheaters », c'est-à-dire des bactéries ayant perdu la capacité à fixer. Ces cheaters peuvent bénéficier de la niche écologique privilégiée que constitue le nodule, dans laquelle ils peuvent proliférer sans (ou avec très peu de) compétiteurs et sans produire de bien en échange. A noter que les rhizobia qui développent une différenciation non terminale bénéficient encore davantage de cette possibilité.

Une théorie veut que pour contrer ce risque de rupture du mutualisme, les plantes aient mis en place des sanctions (co-évolution antagoniste) visant à limiter la prolifération des bactéries Fix-. Les premiers exemples concrets d'existence de sanctions ont été démontrés en mesurant le nombre de rhizobia sur des plantes entières, en split root et sur des nodules individuels dans lesquels les rhizobia ne pouvaient pas fixer puisque l'azote atmosphérique était remplacé par de l'argon (Kiers et al. 2003; Oono et al. 2011). Ces travaux ont montré qu'avec le traitement à l'argon, les nodules étaient plus petits et contenaient moins de bactéries comparés aux nodules contrôles qui avaient accès à l'azote et ce, dans des nodules déterminés ou des nodules indéterminés avec des bactéroïdes différenciés de façon terminale (Kiers et al. 2003; Oono et al. 2011). Le mécanisme mis en place serait une privation de O₂, mais il n'a pas été démontré (Kiers et al. 2003).

Une autre théorie propose que les plantes favorisent les Fix+ en leur allouant des ressources (co-évolution mutualiste) (Frederickson 2013). L'allocation des composés carbonés issus de la photosynthèse aux racines qui apportent le plus d'azote est un trait ancien probablement présent avant l'apparition du mutualisme (Frederickson 2013). Ainsi, la capacité d'allocation des ressources carbonées aurait été sélectionnée chez les légumineuses pour le maintien du mutualisme avec les rhizobia. Cette hypothèse est confortée par certaines études qui ont montrées que la plante récompense ses symbiotes les plus efficaces (Laguerre et al. 2012). C'est pourquoi il n'est pas clair si les réponses observées par les travaux de Kiers et al. sont réellement dues à des sanctions ou plutôt le résultat d'une réallocation sélective du carbone en réponse à l'acquisition de l'azote (Frederickson 2013).

1.3. Évolution expérimentale de symbiotes de légumineuses

Les études phylogénétiques, génomiques et génétiques ont conduit à proposer que les rhizobia aient évolué en deux étapes : acquisition des gènes symbiotiques par transfert latéral et optimisation du potentiel symbiotique acquis par remodelage du génome d'accueil sous pression de sélection des légumineuses hôtes. Dans l'objectif de comprendre les mécanismes moléculaires et évolutifs permettant à une bactérie du sol de devenir un symbiote de

légumineuse, l'équipe de Catherine Masson a initié il y a une dizaine d'années une expérience visant à reproduire en laboratoire l'évolution d'un nouveau genre de rhizobium. Pour cela le β -rhizobium *C. taiwanensis* a été choisi comme donneur des gènes *nod* et la bactérie phytopathogène *Ralstonia solanacearum* comme receveur des gènes symbiotiques. La bactérie résultante a été évoluée sous pression de sélection de *Mimosa pudica*, l'hôte de *C. taiwanensis* en utilisant une approche d'évolution expérimentale. La distance phylogénétique existant entre ces deux bactéries était idéale pour une expérience d'évolution, suffisamment grande pour laisser de la place à l'évolution, suffisamment petite pour que l'évolution se fasse à l'échelle du laboratoire. De plus le choix d'un pathogène permettait d'étudier les transitions pathogénie-symbiose.

I.3.1. L'évolution expérimentale

L'évolution expérimentale est définie comme la propagation d'un organisme vivant dans un environnement contrôlé, tel que le laboratoire ou le champ, sur des centaines voire des milliers de génération (Elena and Lenski 2003). Suivre et analyser l'évolution en action permet de tester des hypothèses quant à l'évolution des organismes (Buckling et al. 2009). L'évolution expérimentale couplée à du re-séquençage (appelé 'Evolved and resequence' ou E&R) permet d'analyser comment les génomes évoluent en réponse aux changements de l'environnement (Brockhurst et al. 2011; Schlötterer et al. 2015). L'E&R est devenu un outil fondamental de la biologie évolutive contemporaine (Schlötterer et al. 2015).

L'évolution expérimentale a été appliquée à de nombreux organismes comme les nématodes (Teotónio et al. 2017), la drosophile (Zhou et al. 2011) ou même la souris (Kelly et al. 2006) mais aussi des virus (Deardorff et al. 2011; Elena et al. 2008). Néanmoins, les bactéries restent les organismes les plus appropriées pour l'évolution expérimentale grâce à la plasticité de leur génome et à leur taux de génération élevé, et également parce qu'elles sont génétiquement manipulables et peuvent être congelées, ce qui permet de conserver les ancêtres et intermédiaires comme « fossiles vivants » (Buckling et al. 2009).

La première évolution expérimentale date de la fin du 19^{ème} siècle, où William Dallinger a été le premier à soumettre des organismes à une augmentation progressive de température (passant de 15°C à 70°C) pendant sept ans. Dallinger a montré que les populations évoluées pouvaient survivre à 70°C mais ne pouvaient plus croître à 15°C, température de croissance au départ de l'expérience (Dallinger 1878). L'expérience d'évolution la plus célèbre, menée par Richard Lenski, suit le développement évolutif de douze populations de *E. coli* depuis plus de 30 ans. Au cours du temps, une grande diversité de changements phénotypiques et génotypiques ont été rapportés sur les populations évoluées, tels que la diminution du fitness, la modification de la taille des cellules ou plus surprenant, l'acquisition de croissance aérobie sur citrate (Lenski et al. 1991; Elena and Lenski 2003; Barrick et al. 2009). L'évolution expérimentale a contribué à la connaissance de la biologie des organismes, en permettant une meilleure compréhension de la régulation des gènes, de la résistance aux antibiotiques et

interactions hôte/microbiome (Remigi et al. 2019). Lorsque les conditions d'évolution expérimentale se rapprochent des conditions naturelles, elles présentent des similitudes avec les processus naturels (Clerissi et al. 2018).

I.3.2. Le modèle *Ralstonia solanacearum*/*Cupriavidus taiwanensis* /*Mimosa pudica*

I.3.2.1 *Cupriavidus taiwanensis*

C. taiwanensis est un rhizobium appartenant aux β -protéobactérie et est, avec *Cupriavidus necator* et *pinatubonensis* (Florentino et al. 2012), une des espèces du genre *Cupriavidus* décrite comme symbiote de légumineuse. Les bactéries du genre *Cupriavidus*, précédemment connu comme *Wausteria* ou même *Ralstonia* (Vandamme and Coenye 2004), sont des bactéries gram négatives, mobiles, en forme de bâtonnet, non sporulées et ayant un métabolisme oxydatif (Chen et al. 2003, 2001). Ce genre contient des espèces provenant d'environnements divers, tels que sols, plantes, animaux (Goris et al. 2001; Karafin et al. 2010; Estrada-de los Santos et al. 2012). *C. taiwanensis* regroupe à la fois des isolats de nodules de légumineuses (*Mimosa*) et des souches cliniques isolées de patients atteints de mucoviscidose (Chen et al. 2001).

C. taiwanensis est présente sur plusieurs continents, l'Amérique (Mishra et al. 2012), l'Océanie avec des souches isolées en Nouvelle-Calédonie (Klonowska et al. 2012), et l'Asie (Verma et al. 2004; Liu et al. 2011). *C. taiwanensis* a été isolé pour la première fois à Taiwan où elle est l'espèce symbiotique dominante des *Mimosa* (Chen et al. 2001, 2003). C'est une espèce symbiotique qui possède un spectre d'hôte étroit puisqu'il semble être confiné aux espèces *M. diplotricha*, *M. pigra* et *M. pudica* (Chen et al. 2001, 2003, 2005) et plus précisément à *M. pudica* var. *unijuga* avec laquelle *C. taiwanensis* est le rhizobium le plus compétitif pour la nodulation (Melkonian et al. 2014). Ce spectre d'hôte restreint peut être élargi à *Leucaena leucocephala* lorsque le SST3 de *C. taiwanensis* est désactivé (Saad et al. 2012). Le SST3 de *C. taiwanensis* est retrouvé chez toutes les souches symbiotiques (Clerissi et al. 2018), et a rôle dans la symbiose bien qu'il semble être indépendant de la régulation *nod* (Saad et al. 2012).

Le génome de *C. taiwanensis* LMG19424 a été séquencé et annoté de façon experte (Amadou et al. 2008). Cette souche se compose de deux chromosomes de 3,42 Mb et 2,50 Mb ainsi que d'un grand plasmide symbiotique de 0,56 Mb (Amadou et al. 2008). Plus récemment, 40 autres souches ont eu leurs génomes séquencés. Ces études ont révélé que *C. taiwanensis* est un rhizobium relativement récent puisque l'acquisition du plasmide symbiotique, pRalta, se serait faite il y a environ 16 millions d'années (Clerissi et al. 2018). C'est également un rhizobium minimal qui possède un cluster de gènes symbiotiques de seulement 35kb, ce qui en fait le plus petit décrit (Amadou et al. 2008). Cet îlot symbiotique comprend 9 gènes *nod*, *nodBCIJHASUQ*, un gène de régulation *nodD* et 19 gènes impliqués dans la fixation de l'azote. Sur le pRalta deux *nod* boxes ont été prédites, la première dans la région intergénique *nodB-*

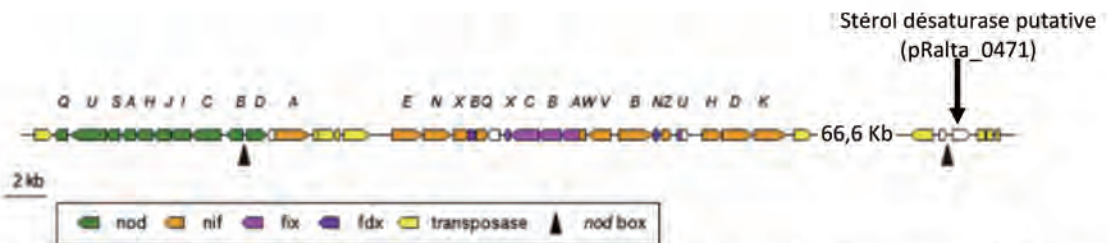


Figure I.3.2.1.1: Région symbiotique de *Cupriavidus taiwanensis* LMG19424. La coloration des gènes dépend de leurs noms. Les pointes indiquent la position des *nod* boxes. La flèche indique la position du gène pRalta_0471. Adapté de (Amadou et al. 2008)

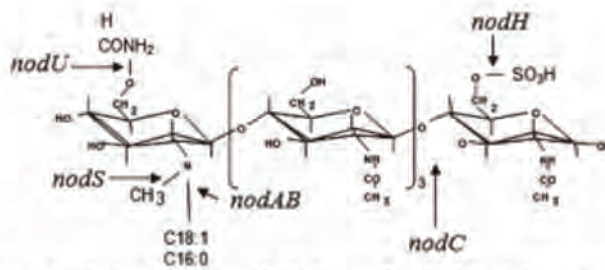


Figure I.3.2.1.2: Structure des facteurs Nod synthétisés par *Cupriavidus taiwanensis* LMG19424. L'analyse MS des FNs produits par *C. taiwanensis* a permis de les décrire comme étant des pentamères de N-acetylglucosamine, sulfatés sur le sucre réducteur et substitués par un groupement acyle (acide palmitique (C_{16:0}) ou un acide vaccénique (C_{18:1})) au niveau de la glucosamine terminale non réductrice. La plupart des résidus portent des substitutions N-méthyl et un groupe carbamyl au niveau du sucre non réducteur (Amadou et al. 2008).

nodD et la seconde localisée en amont d'un gène annoté comme une putative stérol désaturase, *pRalt_0471*, (Amadou et al. 2008) (**Figure I.3.2.1.1**). C'est pour ces caractéristiques, symbiote minimal et récemment évolué que *C. taiwanensis* a été choisi comme donneur des gènes *nod* dans l'expérience d'évolution.

Les FNs synthétisés par la souche type, LMG19424, ont été décrits comme des pentamères de N-acetylglucosamine, sulfatés (*nodH*) sur le sucre réducteur et substitués par un groupement acyle (acide palmitique (C_{16:0}) ou un acide vaccénique (C_{18:1})) au niveau de la glucosamine terminale non réductrice. La plupart des résidus portent des substitutions N-méthyl (*nodS*) et un groupe carbamyl (*nodU*) sur le sucre terminal non-réducteur, ce qui correspond à la prédiction *in silico* (**Figure I.3.2.1.2**). Les gènes de nodulation *nodM* et *nodP* n'ont pas été détectés sur le pSym, suggérant que leurs homologues chromosomique *glmS* et *cysD* participent à la biosynthèse des NFs (Amadou et al. 2008).

L'interaction entre *C. taiwanensis* LMG19424 et *M. pudica* a été décrite : les bactéries pénètrent dans la plante via les poils racinaires de cette dernière dans lesquels les cordons d'infections se développent jusqu'à formation de nodules de type indéterminé avec un méristème persistant, une zone d'invasion et une zone de fixation contenant les symbiosomes (Chen et al. 2003; Marchetti et al. 2011). Les symbiosomes des nodules contiennent jusqu'à 4 bactéroïdes qui sont décrits comme non différenciés de façon terminale puisque 20% de ces bactéroïdes peuvent pousser sur boîte (Marchetti et al. 2011). Cette caractéristique est une différence majeure comparée aux nodules indéterminés formés par les *Aeschynomene* et les plantes du IRLC (Mergaert et al. 2006; Kereszt et al. 2011; Czernic et al. 2015).

I.3.2.2 *Ralstonia solanacearum*

R. solanacearum a été choisi comme receveur des gènes *nod-nif* car le genre *Ralstonia*, voisin du genre *Cupriavidus* dans la classe des β -protéobactéries, ne contient pas de rhizobia. *R. solanacearum* est un pathogène facultatif qui présente deux modes de vie : le mode saprophytique dans le sol, dans lequel la bactérie peut survivre de nombreuses années et le mode infectieux dans la plante (van Elsas et al. 2000). Il est responsable du flétrissement bactérien sur plus de 200 espèces végétales dont de nombreuses avec un intérêt agronomique et économique telles que la tomate, la pomme de terre ou encore le tabac (Hayward 1991). La répartition mondiale de *R. solanacearum* (présent sur tous les continents), son large spectre d'hôte, son impact agronomique et économique néfaste, ainsi que sa capacité de transformation naturelle ont fait de cette bactérie un modèle très étudié.

Le séquençage et l'annotation du génome de *R. solanacearum* GMI1000 en 2002 a permis d'établir le répertoire complet des effecteurs de type 3 et donc d'en faciliter l'étude (Salanoubat et al. 2002). Le génome de GMI1000 se compose d'un chromosome d'environ 3,5 Mb et d'un mégaplasmide de 2 Mb sur lequel sont présents tous les gènes *hrp* (pour 'hypersensitive response and pathogenecity') qui sont requis pour provoquer la maladie chez les plantes en régulant le SST3, entre autres (Salanoubat et al, 2002).

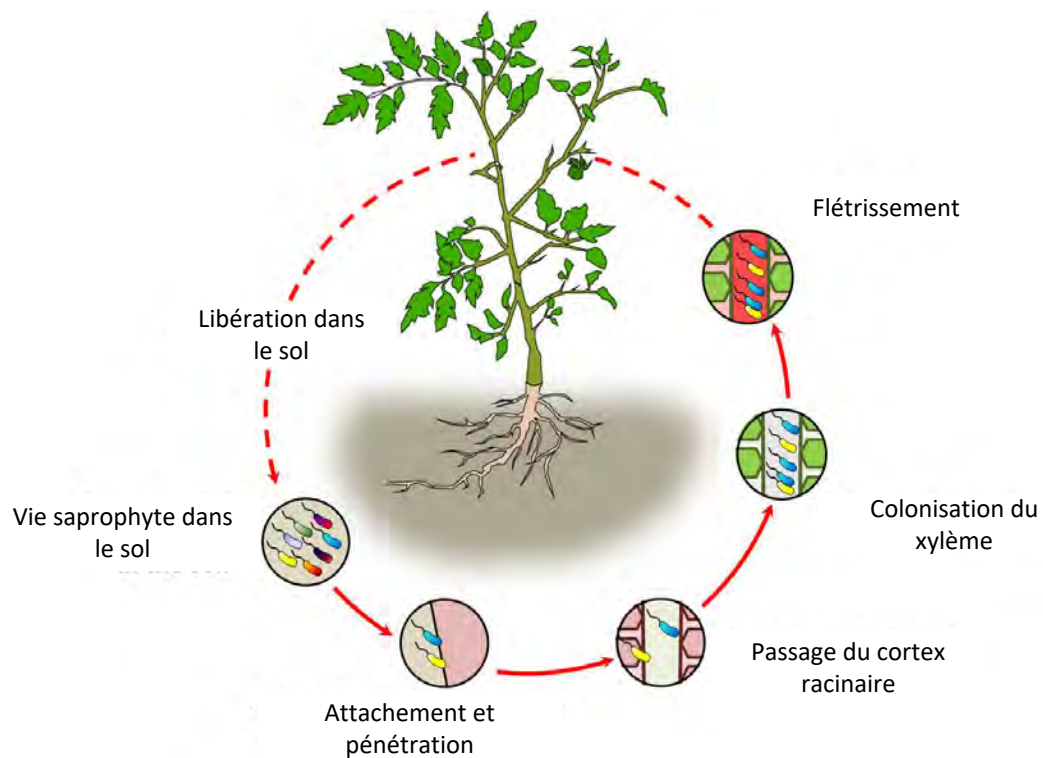


Figure 1.3.2.2.1: Cycle infectieux de *Ralstonia solanacearum*. le mode saprophytique dans le sol *R. solanacearum* pénètre dans ses plantes hôtes via des blessures naturelles ou non des racines de la plante. La bactérie se multiplie dans les espaces interstitielles du cortex racinaire et du parenchyme vasculaire puis envahit les vaisseaux du xylème, permettant ainsi la colonisation systémique de l'hôte. La multiplication bactérienne dans le xylème s'accompagne de production massive d'EPS dont l'accumulation bloque la circulation de la sève brute, provoquant la maladie. A la mort de la plante, les bactéries retournent à une vie saprophyte dans le sol.

R. solanacearum infecte ses plantes hôtes par les racines, au niveau de la coiffe racinaire, des zones d'émergence des nouvelles racines secondaires ou encore par les blessures. La bactérie se multiplie dans les espaces interstitielles du cortex racinaire et du parenchyme vasculaire dont les parois cellulaires sont dégradées puis envahit les vaisseaux du xylème, permettant ainsi la colonisation systémique de l'hôte (Vasse 1995). La multiplication bactérienne dans le xylème atteint jusqu'à 10^{10} CFU/mg de tissu de tige de tomate (Grimault et al. 1994) et s'accompagne de la production massive d'EPS dont l'accumulation bloque la circulation de la sève brute, provoquant la maladie (Grimault et al. 1994; Genin and Denny 2012) (**Figure I.3.2.2.1**). Chez *R. solanacearum* la virulence, ou capacité de la bactérie à se développer dans la plante, et la pathogénicité, ou capacité à provoquer la maladie, sont régulées de façon sophistiquée par des réseaux interconnectés. A chaque stade de l'infection, ces caractères sont régulés en réponse à des conditions environnementales spécifiques, telles que la densité bactérienne ou des facteurs végétaux (Genin and Denny 2012).

De nombreux facteurs de virulence/pathogénicité ont été décrits : flagelles, enzymes de dégradation des parois végétales et systèmes de sécrétion. Les flagelles sont impliqués dans les étapes précoces de l'infection, en permettant à la bactérie de se déplacer jusqu'à l'entrée de la racine hôte puis en rendant possible la colonisation des vaisseaux du xylème où la bactérie deviendra alors essentiellement non mobile (Tans-Kersten et al. 2004). Le passage de la bactérie dans le xylème est rendu possible par la dégradation des parois cellulaires du parenchyme par des exoenzymes, appelées aussi enzymes extracellulaires. Ces enzymes sont excrétées par l'un des systèmes de sécrétion de *R. solanacearum*, qui en compte plusieurs. Les SST2 et SST3 sont les deux systèmes de sécrétion les plus impliqués dans la virulence. Le SST2 permet la sécrétion des exoenzymes impliquées dans la dégradation des tissus racinaires. Le SST3 est quant à lui l'un des déterminants majeurs au pouvoir pathogène de la bactérie puisque c'est l'acteur essentiel de l'infection des plantes par la bactérie (Arlat et al. 1992; Boucher et al. 1985). Véritable seringue moléculaire encodée par les clusters de gènes *hrp* et *hrc* (*hrp* conserved), elle permet d'injecter des effecteurs de types III (T3Es) dans la cellule cible de l'hôte (Valls et al. 2006; Genin et al. 2005). Chaque pathogène possède un jeu d'effecteurs qui lui est propre. Chez la souche GM11000 74 effecteurs de types III ont été recensés (Poueymiro and Genin 2009). La sécrétion dans le milieu extracellulaire ainsi que la translocation dans la cellule végétale a été démontré pour 28 d'entre eux (Poueymiro and Genin 2009). La sécrétion des T3Es dans la plante par la bactérie permet la suppression de l'immunité chez l'hôte et est responsable de la réponse hypersensible chez les plantes non hôtes (Coll and Valls 2013; Arlat et al. 1992; Boucher et al. 1985). Les EPS sont aussi des facteurs de virulence importants de *R. solanacearum* puisque des mutants déficients pour la production d'EPS sont quasiment avirulents (Denny and Bayek, 1991; Kao et al, 1992). C'est l'accumulation des EPS dans les vaisseaux du xylème qui conduit à leur occlusion durant les stades tardifs de l'infection.

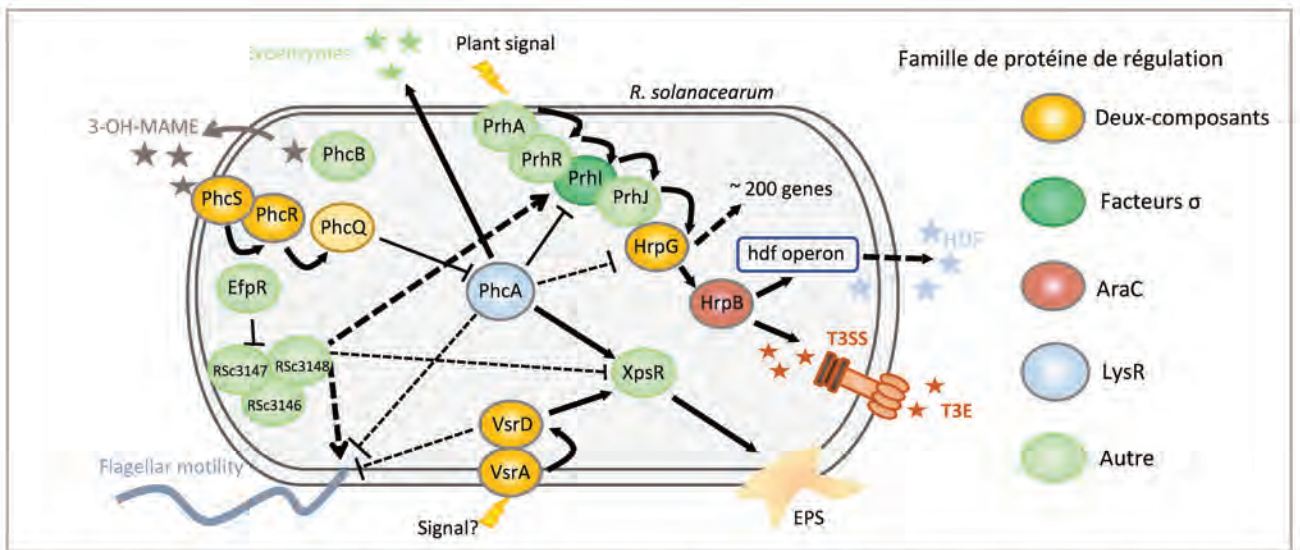


Figure I.3.2.2.2: Principales voies de régulation de la virulence de *Ralstonia solanacearum*. Régulation des gènes et traits directement régulés positivement ou négativement par les régulateurs de transcription centraux. (Tang et al. 2020).

Plusieurs régulateurs centraux, tels que PhcA, VsrA, HrpG ou encore EfpR ont été identifiés pour être impliqués dans le contrôle de l'expression des facteurs de virulence. Ces régulateurs sont interconnectés et contrôlés de façons différentes (Genin and Denny 2012).

PhcA est un régulateur central de type LysR qui régule de façon directe ou indirecte l'expression de nombreux facteurs de virulence tels que la production d'EPS et d'exoenzymes, la motilité bactérienne et le SST3 (Genin and Denny 2012; Genin et al. 2005). Il permet par exemple, avec *vsrA*, d'activer l'expression du gène *xpsR* et ainsi permettre la sécrétion d'EPS. C'est aussi PhcA qui active l'expression des gènes *egl*, *pme* et *cbha* codant pour les exoenzymes sécrétées par le SST2. PhcA est impliqué dans la répression de plusieurs gènes dont *hrpG* et *pehSR* régulant ainsi indirectement l'expression des gènes du SST3 et de mobilité, respectivement (Genin et al. 2005; Genin and Denny 2012). Il a été démontré que le régulateur PhcA est aussi impliqué dans la répression de voies métaboliques puisque qu'un mutant *phcA* est capable de métaboliser 17 substrats supplémentaires comparé à la souche sauvage (Peyraud et al. 2016). L'activité de PhcA est auto-induite via un système de quorum sensing codé par l'opéron *phcBSRQ*. Cette régulation repose sur l'accumulation de 3-OH MAME (Methyl 3-hydroxymyristate) produit par une méthyltransférase, *phcB*, et reconnu par un système à deux composants PhcS-PhcR, permettant l'augmentation de PhcA fonctionnelle (Clough et al. 1997). Le système Phc (pour phenotypic conversion) est donc au cœur du réseau de régulation du pouvoir pathogène de *R. solanacearum* et est interconnecté aux autres voies de régulation, notamment avec VsrA et HrpG (**Figure I.3.2.2.2**).

La protéine transmembranaire VsrA forme avec VsrD un système de régulation à deux composants qui contrôle de nombreux traits de virulence tels que l'habilité à coloniser et à se multiplier dans l'hôte ou encore la motilité. Le complexe VsrA-VsrD est aussi impliqué dans la production de métabolites secondaires tels que la ralfuranone, dont le rôle dans la pathogénicité n'a pas été démontrée (Genin and Denny 2012). VsrA-VsrD sont principalement connus pour leur rôle dans la régulation positive du gène *xpsR* (production des EPS) en relation avec PhcA. En réprimant *flhDC*, VsrA-VsrD réprime de façon indirecte les gènes *fli* et empêche donc la motilité de la bactérie (**Figure I.3.2.2.2**).

HrpG est une protéine de régulation appartenant à la famille des OmpR (Brito et al. 1999) impliquée dans la régulation positive de nombreux gènes dont le régulateur *hrpB*, de la famille des AraC (Brito et al. 1999). HrpG contrôle l'expression des fonctions qui permettent à la bactérie de s'adapter à la vie dans l'hôte (Genin 2010). HrpG permet aussi l'expression de gènes codants pour certains facteurs de virulence, tel que le SST3 (seringue et effecteurs) dont la régulation se fait via le régulateur HrpB (Genin and Denny 2012). La régulation par HrpG inclut aussi des phytohormones, des enzymes métaboliques, probablement des régulateurs intermédiaires et de nombreux autres gènes de fonctions inconnues (Genin and Denny 2012). HrpB et HrpG sont des composants résultant de la cascade de régulation qui inclut le récepteur transmembranaire PrhA et les protéines PrhIJ. Aux premières étapes de l'infection PrhA reconnaît la présence de composés (de nature encore inconnue) de la paroi végétale et active la cascade de régulation Prh qui aboutit à l'activation de HrpG et HrpB, et donc la mise en



Figure I.3.2.2.2.1: Photo d'un plant de *Mimosa pudica*.

place du SST3 (Aldon et al. 2000). HrpB est aussi responsable de la production d'un auto-inducteur, l'HDF ('HrpB Dependant diffusible Factor'), identifié comme une 3-hydroxy-oxindole (Delaspre et al. 2007). Cette HDF est impliqué dans la perception de la densité cellulaire dans les étapes précoces de l'infection (**Figure I.3.2.2.2**).

EfpR a récemment été démontré pour être un répresseur catabolique clef dans la virulence de *R. solanacearum* (Perrier et al. 2016; Capela et al. 2017). Environ 200 gènes sont régulés par EfpR, dont 124 sont activés et 76 sont réprimés (Capela et al. 2017). La plupart des gènes régulés par EfpR sont également contrôlés par PhcA tel que *xpsR* qui est fortement réprimé mais aussi les gènes *hrp* qui, *a contrario*, sont régulés positivement par EfpR. Un cluster de gènes, *RSc3146-RSc3148*, a été identifié pour être réprimé par EfpR (Capela et al. 2017). Ces gènes de fonctions inconnues sont des acteurs intermédiaires majeurs de la voie de régulation de EfpR et influent, entre autres, sur la production d'EPS, la motilité et le SST3 (Capela et al. 2017). EfpR semble être fortement impliqué dans les voies métaboliques de la bactérie étant donné que chez un mutant *efpR* 62 activités métaboliques sont activées comparés à la souche sauvage. Peu de choses sont connues quant à la régulation de *efpR* si ce n'est qu'une mutation SNP 115 nucléotides en amont du gène RSc0965 réduit l'expression de *efpR*, suggérant une régulation en amont de ce dernier par un acteur non identifié (Capela et al. 2017) (**Figure I.3.2.2.2**).

I.3.2.3 Mimosa pudica

Mimosa pudica (*M. pudica*), communément appelé sensitive, est une plante épineuse à fleurs du clade des Eudicotylédone et appartenant à la famille des *Fabaceae* (**Figure I.3.2.2.1**) la sous famille des *Caesalpinioideae* et la tribu des *Mimoseae* (Magallon et al. 1999; Simon et al. 2011a). C'est une plante tétraploïde ($2n=52$) dont le génome d'environ 500 Mb a été séquencé mais non assemblé (Griesmann et al. 2018). L'analyse du génome de *M. pudica* a permis la détection des gènes essentiels à la symbiose, à l'exception de NAP1 et RPG qui seraient fragmentés (Griesmann et al. 2018). *M. pudica*, originaire d'Amérique du Sud, est une espèce invasive en Asie et en Océanie (Simon et al. 2011b; Chen et al. 2001; Gehlot et al. 2013). C'est une légumineuse avec une forte affinité pour les β -rhizobia, *Burkholderia* et *Cupriavidus*, qui est capable de s'adapter à des milieux extrêmes, comme les sols ultramafiques de Nouvelle-Calédonie (Chen et al. 2003; Bontemps et al. 2010; Melkonian et al. 2014; Klonowska et al. 2012). Une méthode de transformation de *M. pudica* par *Agrobacterium tumefaciens* a été développée par une équipe japonaise afin d'étudier les gènes de la plante impliqués dans le mouvement de ses feuilles (Mano et al. 2014). La transformation par *A. rhizogenes* a été testée et optimisée dans l'équipe. Bien que *M. pudica* ne soit pas une plante modèle, la disponibilité de sa séquence ainsi que les outils de transformation permettent d'envisager d'analyser les réponses de la plante à l'infection par ses partenaires symbiotiques et des bactéries symbiotiquement évoluées.

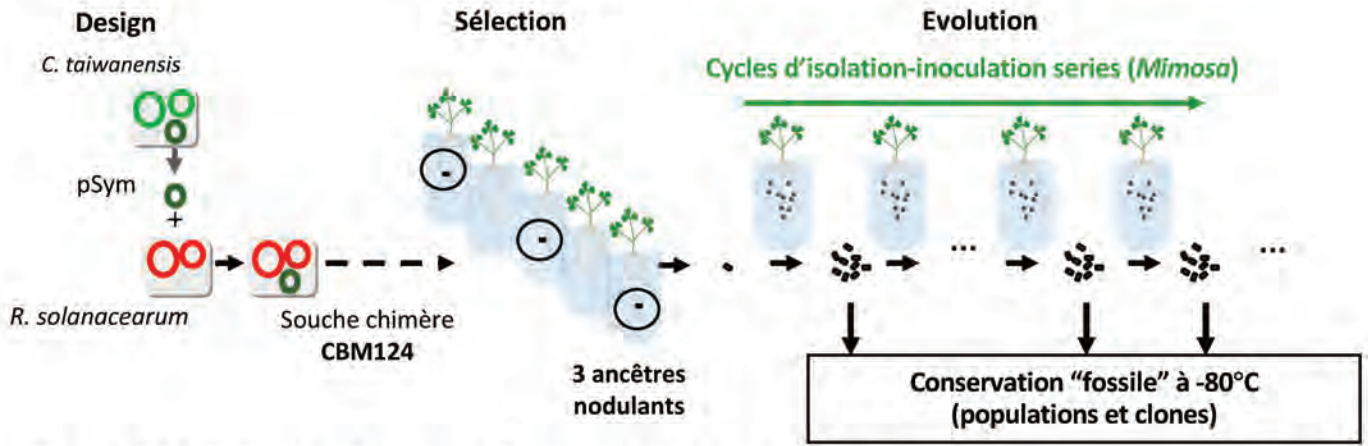


Figure I.3.3.1.1: Design et développement de l'expérience d'évolution. Modèle d'évolution en deux étapes avec transfert des gènes *nod-nif* suivi de l'optimisation du génome d'accueil sous pression de sélection de la plante.

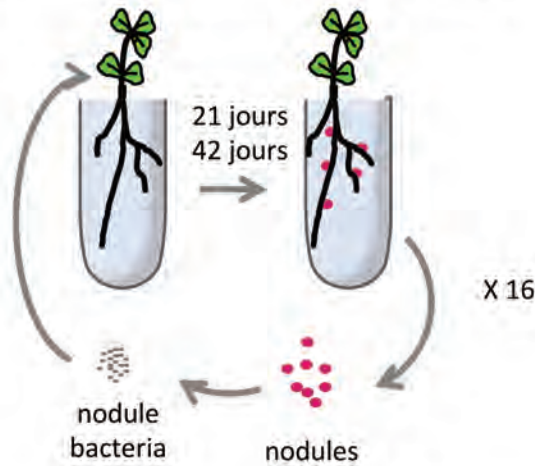


Figure I.3.3.1.2: Alternance de la vie saprophytique et symbiotique en laboratoire.

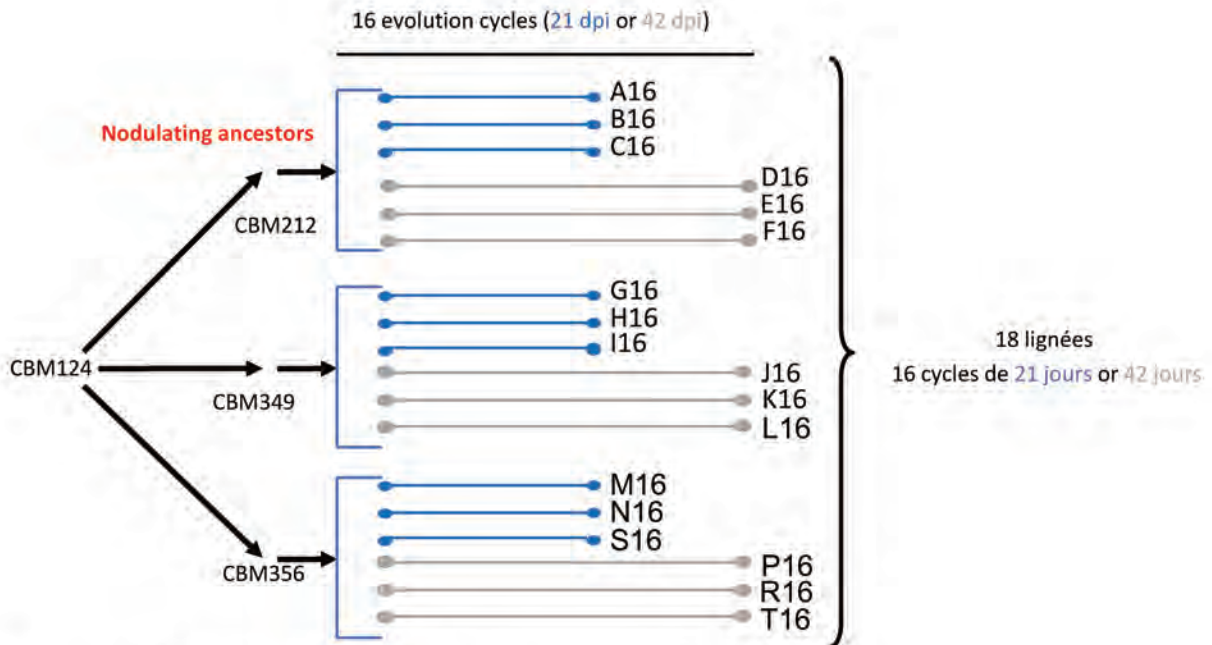


Figure I.3.3.1.3: Régimes de sélection des lignées.

I.3.3. L'expérience d'évolution

I.3.3.1. Design de l'expérience

L'expérience initiée dans le groupe visait à reproduire l'évolution d'un nouveau genre de rhizobium en laboratoire, via le transfert des gènes symbiotiques dans une bactérie du sol suivi de l'optimisation du potentiel symbiotique acquis sous pression de sélection d'une légumineuse. L'étape d'optimisation a été réalisée en reproduisant l'alternance de vie saprophyte (en dehors du nodule) et symbiotique (dans le nodule) qui a façonné l'évolution naturelle des rhizobia (**Figure I.3.3.1.1**).

Ainsi, le plasmide symbiotique de *C. taiwanensis* LMG19424, pRalta, portant les gènes essentiels à la nodulation et à la fixation de l'azote, a été transféré chez *R. solanacearum* GM11000 générant une bactérie chimère toujours pathogène et non nodulante. Après inoculation massive de *M. pudica*, trois nodules sont apparus à partir desquels trois mutants ont été isolés, CBM212, CBM349 et CBM356. Ces 3 bactéries présentaient des capacités d'infection différentes, CBM212 et CBM349 étant capable d'infecter quelques cellules du nodule tandis que CBM356 était exclusivement intercellulaire. Ces ancêtres nodulants ont ensuite été évolués par des cycles d'inoculation/ré-isolation des nodules (populations isolées des nodules stérilisés en surface et directement ré-inoculées à un nouveau set de plantes) (**Figure I.3.3.1.2**). Deux régimes de sélections ont été utilisés : 21 et 42 jours sur 18 lignées évoluées en parallèle, le premier régime ayant pour objectif d'optimiser la nodulation et le second la persistance dans les nodules (**Figure I.3.3.1.3**). A chaque cycle, les populations issues des nodules ainsi qu'un clone isolé par population ont été conservées à -80°C pour des analyses phénotypiques et génotypiques ultérieures.

I.3.3.2. Analyse des mécanismes moléculaires et évolutifs de l'adaptation

L'analyse de l'évolution phénotypique de clones isolés tout au long de l'expérience a permis de valider partiellement l'hypothèse de départ, à savoir celle d'un scénario évolutif en deux étapes (Marchetti et al. 2010, 2014).

Le phénotypage des clones finaux issus des lignées dérivant des ancêtres CBM212 et CBM349 a permis de mettre en évidence une forte évolution parallèle des traits symbiotiques tels que la compétition pour la nodulation, l'infection intracellulaire ou encore la persistance des bactéroïdes (Marchetti et al. 2014). De façon simultanée, une diminution des zones nécrosées des nodules ont été observées indiquant que les réactions de défense de la plante ont diminué (Marchetti et al. 2014). Par contre, pour les lignées dérivant de l'ancêtre intercellulaire CBM356, seuls les cycles courts ont permis l'acquisition stable de l'infection intracellulaire. Les lignées évoluées avec des cycles de 42 jours ont amélioré leur capacité à infecter les nodules de façon intercellulaire (Marchetti et al. 2016).

L'analyse des génotypes des bactéries évoluées a révélé que la reconfiguration des circuits de régulation du génome d'accueil a été le principal moteur d'évolution. Le re-séquençage des

clones ancestraux CBM212, CBM249 et CBM356 et de leurs clones évolués a permis de détecter les mutations acquises (indels, SNPs (polymorphisme d'un seul nucléotide) et délétions) et d'identifier parmi celles-ci, les mutations adaptatives pour l'acquisition et/ou l'amélioration des traits symbiotiques (Marchetti *et al.*, 2010, 2014, 2016; Guan *et al.*, 2013; Tang *et al.*, en préparation).

La nodulation a été acquise chez les ancêtres par inactivation du SST3. Chez CBM356 grâce à une mutation stop dans le gène *hrcV*, mais cette mutation adaptative pour la nodulation ne permet pas l'acquisition de l'infection intracellulaire (Marchetti *et al.* 2010). Chez les clones CBM212 et CBM249, des mutations stop dans le gène *hrpG* permettent la nodulation et induisent un premier niveau d'infection intracellulaire (Marchetti *et al.* 2010). Dans les lignées de cycles courts M, N and S dérivant de CBM356, des mutations dans les gènes *vsrA* et *prhI* permettent un premier niveau d'infection intracellulaire (Guan *et al.* 2013).

L'infection intracellulaire a été améliorée dans les lignées de CBM212 et CBM249. Des mutations dans les gènes *efpR* et *phcB* ont été identifiées comme responsables de cette amélioration de l'infection intracellulaire (Capela *et al.* 2017).

Cette expérience a également permis la découverte d'une cassette de mutagenèse *imuABC* présente sur le plasmide symbiotique de *C. taiwanensis*. Cette cassette déclenche une hypermutabilité transitoire chez les clones évolués de *R. solanacearum* dans le milieu de culture ou dans la rhizosphère qui induit l'accélération de l'évolution symbiotique de GMI1000 en symbiote de *M. pudica* sous pression de sélection de cette dernière (Remigi *et al.* 2014). De façon intéressante, 50% des plasmides symbiotiques possèdent une cassette *imuABC* suggérant que ce mécanisme opère dans la nature (Remigi *et al.* 2014).

De plus, des parallèles entre l'évolution expérimentale et naturelle de symbiotes de *M. pudica* ont été identifiés, ce que ne laissait pas obligatoirement présager les différences entre les deux processus en termes d'ancêtres, de timing, issue et conditions de l'évolution (Clerissi *et al.* 2018).

I.3.4. Projet de thèse

A ce stade de l'expérience trois observations ont été à l'origine de mon projet de thèse : i) les bactéries évoluées ne fixent pas l'azote après 16 cycles et l'évolution vers le mutualisme n'est donc pas achevée à ce stade, ii) un gène de fonction inconnue semble important pour l'infection intracellulaire et iii) les mutations permettant l'acquisition et/ou l'amélioration de l'infection des cellules du nodule semblent également améliorer la nodulation.

L'absence de fixation à la fin des 16 cycles de l'expérience d'évolution pourrait être dû à un régime de sélection ne permettant pas à un mutant Fix+ d'être sélectionné dans l'expérience. Le premier objectif de ma thèse a donc été d'évaluer les conditions (en particulier longueur des cycles) qui permettraient l'augmentation en fréquence et la fixation d'une sous-population fixatrice (Fix+) au sein d'une population non fixatrice (Fix-) au cours de cycles

répétés d'inoculation sur *Mimosa pudica*. Cela m'a amené à analyser la dynamique de populations quasi-isogéniques de *C. taiwanensis* Fix+ et Fix- au cours du processus infectieux ou d'infections répétées.

Dans la lignée extracellulaire R, la présence d'une mutation adaptative conférant l'intracellularité a été identifiée. Les travaux sur cette lignée ont montrés que des mutations sur le pR_{Alta} contrebalanceraient l'effet de cette mutation. La délétion du gène *pR_{Alta}_0471* est apparue comme responsable du maintien de l'infection extracellulaire des nodules dans cette lignée. Mon deuxième objectif a donc été de déterminer le rôle de ce gène dans le rhizobium de référence, *C. taiwanensis*.

Les mutations adaptatives clefs pour l'infection, ciblant les gènes *hrpG*, *efpR*, *Rsc0965*, *phcQB*, avaient également un impact positif sur la nodulation, suggérant un lien entre infection et nodulation. Le troisième objectif de ma thèse a donc été de réaliser, une étude cytologique détaillée de l'impact symbiotique de ces mutations au cours des étapes précoces de l'infection.

Chapitre II : Le contrôle spatio-temporel du mutualisme chez les légumineuses favorise la dissémination de la symbiose fixatrice d'azote

II.1. Introduction

L'expérience d'évolution de *R. solanacearum* en symbiotes de *M. pudica* a rapidement permis d'acquérir la nodulation et l'amélioration des capacités infectieuses dans de nombreuses lignées. Cependant, au terme des 16 premiers cycles, aucune lignée n'était capable de fixer l'azote au bénéfice de la plante, en symbiose.

Plusieurs hypothèses ont été faites pour expliquer l'absence de fixation :

- i) Il est vraisemblable que l'évolution d'un pathogène en symbiote efficace soit un processus nécessitant de nombreuses mutations et donc soit suffisamment lent pour ne pas être achevé en 16 cycles et 400 générations seulement.
- ii) L'absence de fixation pourrait être due à des mutations délétères dans les gènes *nif/fix*. Cependant cette seule raison semble invalidée par le fait que le remplacement du pR_{Alta} évolué par l'original dans plusieurs clones finaux n'ait pas modifié le caractère fixateur.
- iii) La fixation pourrait être impossible du fait de gènes manquants dans le génome de *R. solanacearum*.
- iv) Enfin il est possible que les régimes de sélection utilisés ne soient pas favorables. Dans le cas où des mutations permettant un certain niveau de fixation d'azote apparaîtraient, la principale interrogation était de savoir si les conditions de l'expérience d'évolution permettraient leur augmentation en fréquence au sein d'une population majoritairement non fixatrice.

Le régime de sélection étant un facteur sur lequel nous pouvions facilement jouer, nous avons évalué les conditions favorables à la montée en fréquence d'un allèle Fix⁺. Pour cela, nous sommes revenu au système de référence *C. taiwanensis*/*M. pudica* pour lequel nous avons une souche Fix⁺ et une souche Fix⁻. Dans un premier temps nous avons analysé la dynamique de populations mixtes de Fix⁺ / Fix⁻ au cours du processus infectieux, ce qui nous a permis de mettre en évidence un contrôle spatio-temporel des Fix⁻ par la plante, y compris au sein d'un même nodule. Dans un second temps, un modèle permettant de simuler la dynamique des populations bactériennes au cours des cycles de l'évolution a été développé. Le modèle a pu être paramétré à partir des données expérimentales et a permis de faire des simulations permettant d'explorer l'effet de différents paramètres écologiques sur la montée en fréquence d'une sous population Fix⁺ dans une population majoritairement Fix⁻. Nous avons ensuite pu valider ce modèle expérimentalement dans deux conditions. Ces travaux sont

décrits dans l'article *Daubech et al., 2017, Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation*, publié dans la revue *eLife*.

II.2. Article

Ma contribution dans ce travail a été d'étudier le devenir des bactéries Fix+ et Fix- au cours du processus infectieux ou après plusieurs cycles de nodulation. J'ai également préparé les échantillons pour la microscopie électronique et réalisé les observations cytologiques. Enfin, j'ai contribué à la relecture de la publication ainsi qu'à la production des figures.

Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation

Benoit Daubech^{1†}, Philippe Remigi^{2†}, Ginaini Doin de Moura¹, Marta Marchetti¹, Cécile Pouzet³, Marie-Christine Auriac^{1,3}, Chaitanya S Gokhale⁴, Catherine Masson-Boivin^{1*}, Delphine Capela¹

¹The Laboratory of Plant-Microbe Interactions, Université de Toulouse, INRA, CNRS, Castanet-Tolosan, France; ²New Zealand Institute for Advanced Study, Massey University, Auckland, New Zealand; ³Fédération de Recherches Agrobiosciences, Interactions et Biodiversité, Plateforme d'Imagerie TRI, CNRS - UPS, Castanet-Tolosan, France; ⁴Research Group for Theoretical Models of Eco-evolutionary Dynamics, Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, Plön, Germany

Abstract Mutualism is of fundamental importance in ecosystems. Which factors help to keep the relationship mutually beneficial and evolutionarily successful is a central question. We addressed this issue for one of the most significant mutualistic interactions on Earth, which associates plants of the leguminosae family and hundreds of nitrogen (N₂)-fixing bacterial species. Here we analyze the spatio-temporal dynamics of fixers and non-fixers along the symbiotic process in the *Cupriavidus taiwanensis*–*Mimosa pudica* system. N₂-fixing symbionts progressively outcompete isogenic non-fixers within root nodules, where N₂-fixation occurs, even when they share the same nodule. Numerical simulations, supported by experimental validation, predict that rare fixers will invade a population dominated by non-fixing bacteria during serial nodulation cycles with a probability that is function of initial inoculum, plant population size and nodulation cycle length. Our findings provide insights into the selective forces and ecological factors that may have driven the spread of the N₂-fixation mutualistic trait.

DOI: <https://doi.org/10.7554/eLife.28683.001>

*For correspondence:

catherine.masson@inra.fr

†These authors contributed equally to this work

Competing interests: The authors declare that no competing interests exist.

Funding: See page 17

Received: 17 May 2017

Accepted: 11 October 2017

Published: 12 October 2017

Reviewing editor: Wenying Shou, Fred Hutchinson Cancer Research Center, United States

© Copyright Daubech et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

The evolutionary dynamics of mutualistic interactions between higher organisms and microbes depends to a large extent on the transmission mode of microbial symbionts. Vertical transmission is expected to promote fitness alignment of obligate symbionts and their partners (*Herre et al., 1999*). In contrast, horizontal transmission generates more complex ecological cycles for facultative symbionts. When going through these cycles, microbes are subjected to several trade-offs regarding host range (specialist vs. generalist) and investment in the mutualism (good or bad cooperators, life in the host vs. outside the host). The large number of possible strategies to maximize fitness, and the ability to segregate in a population of genetically variable partners, often entails conflicts of interests between symbionts and their hosts (*Bever et al., 2009; Sachs et al., 2010; Porter and Simms, 2014; Jones et al., 2015*) that may result in the classic Tragedy of the Commons (*Hardin, 1968*). The emergence and stability of mutualism thus requires that proliferation of symbionts is allowed but restricted to appropriate spaces and times and that beneficial partners are ultimately favored over uncooperative ones (*Vigneron et al., 2014; Visick and McFall-Ngai, 2000; Koch et al., 2014*). The theoretical aspects of the evolution and maintenance of mutualistic

eLife digest Rhizobia are soil bacteria that are able to form a symbiotic relationship with legumes – plants that include peas, beans and lentils. The bacteria move into cells in the roots of the plant and cause new organs called nodules to form. Inside the nodules the bacteria multiply before being released to the soil again. Also while in the nodules, the bacteria receive carbon-containing compounds from the plant. In return many of the bacteria convert (or “fix”) nitrogen from the air into compounds that the plant can use to build molecules such as DNA and proteins. Yet, some of the bacteria are “non-fixers” that provide little or no benefit to the host plant.

Evidence suggests that legumes select against non-fixer bacteria, though it was not clear when or how this selection process occurs. Daubech, Remigi et al. have now followed the number and viability of two variants of a bacteria species called *Cupriavidus taiwanensis* as they form a symbiotic interaction with *Mimosa pudica*, a member of the pea family. The two types of bacteria differed only by whether or not they were able to fix nitrogen. At first fixers and non-fixers entered nodules and multiplied at equal rates. Later, the fixers progressively outcompeted the non-fixers. Then, around 20 days after the bacteria entered the plant, nodule cells that contained non-fixers degenerated. This indicates that the nodule cells help to control bacterial proliferation based on the benefits they receive in return.

Further experiments and mathematical modeling also showed that over repeated cycles of root nodule formation, nitrogen fixers can invade a bacterial population dominated by non-fixer bacteria. The likelihood that this invasion will be successful increases as three other factors increase: the proportion of fixer bacteria in the initial population, the number of available plants, and the length of time the bacteria spend in the nodules. This mechanism ensures the maintenance and spread of nitrogen-fixing traits in the bacterial population.

Improving the processes of biological nitrogen fixation could help to reduce the amount of fertilizers required to grow crops. This in the future could help make agricultural ecosystems more sustainable. The results presented by Daubech, Remigi et al. provide guidelines that could be used to select nitrogen-fixing bacteria on legume crops or on nitrogen-fixing cereals that may be engineered in the future. Further work is now needed to understand in more detail the molecular mechanisms that lead to the death of non-fixer bacteria.

DOI: <https://doi.org/10.7554/eLife.28683.002>

interactions have been extensively discussed (**Archetti et al., 2011; Akcay, 2015**). Yet experimental assessment is scarce and the impact of ecological factors, such as population size of hosts and symbionts or the duration of the interaction, has been under-explored, although they are an essential component of the evolutionary potential of symbiotic systems.

Rhizobia, the N₂-fixing symbionts of legumes, induce the formation of and massively colonize nodules, where intracellular bacteria fix atmospheric nitrogen for the benefit of the plant in exchange for photosynthates. When the nodule senesces, nodule bacteria are released to the soil where they can return to free-living lifestyle and/or colonize a new host (**Thies et al., 1995**). During evolution, symbiosis modules carrying genes essential for the symbiotic process have spread to many different taxa so that extant rhizobia are distributed in hundreds of species in 14 genera of α - and β -proteobacteria (**Remigi et al., 2016**). Acquisition of symbiotic genes may not be sufficient to create an effective symbiont and may lead to bacteria exhibiting various levels of symbiotic capacities (**Nandasena et al., 2006; Nandasena et al., 2007; Marchetti et al., 2010**) that can be further optimized and maintained under legume selection pressure (**Marchetti et al., 2017; Marchetti et al., 2014**). It has been established that bacteria better able to form and infect nodules are selected by a partner choice mechanism involving the specific recognition of bacterial molecular signals by plant receptors (**Kawaharada et al., 2015; Radutoiu et al., 2003**). Bacterial features that are recognized by the plant include Nod factors that initiate rhizobial entry and nodule formation (**Perret et al., 2000; Broghammer et al., 2012**), and lipo/exopolysaccharides critical for root infection and bacterial release inside the plant cell (**Kawaharada et al., 2015**), as well as an array of bacterial effectors that refine host specificity (**Deakin and Broughton, 2009**). Nitrogen fixation however is uncoupled from nodulation and infection, and legumes can be nodulated and infected by

ineffective symbiotic partners (Gehlot *et al.*, 2013; Gourion *et al.*, 2015). The emergence of mutualism in populations resulting from the transfer of symbiosis modules, and its maintenance over evolutionary timescales (Werner *et al.*, 2014) indicates that the cooperative behaviour of the bacterial symbionts is controlled at the infection and/or post-infection levels by one or a combination of mechanisms. Partner choice is the selection of appropriate symbionts at the (pre-) infection stage based on signal recognition while post-infection sanctions rely on the ability to discriminate between low- and high-quality cooperators during an established interaction and to punish or reward them accordingly (Kiers and Denison, 2008; Frederickson, 2013). Partner-fidelity feedback (PFF) ensures positive assortment of symbionts during long lasting or repeated interactions in spatially structured environments independently from any recognition process or conditional response (Sachs *et al.*, 2004). These different control mechanisms have been proposed to affect the dynamics of mutualistic traits, particularly in the context of the nitrogen-fixing symbiosis (Kiers *et al.*, 2003; Oono *et al.*, 2009). Here we evaluate how selective forces and ecological factors act on the dissemination of the nitrogen fixation mutualistic trait on the *Cupriavidus taiwanensis*-*Mimosa pudica* mutualistic interaction. Specifically we evaluated the spatio-temporal dynamics of N₂-fixing and non-fixing bacterial subpopulations to model the spread of the N₂-fixation trait across plant generations.

Results

Evidence for a spatial and temporal control of mutualism in *Mimosa* nodules

During the symbiotic process, most rhizobia enter the legume root via infection threads that ensure colonization of the forming nodule and ultimately release bacteria into nodule cells where differentiated forms called bacteroids fix nitrogen (Batut *et al.*, 2004). Although they induce the formation of indeterminate nodules, it is noteworthy that *Cupriavidus taiwanensis* symbionts of *Mimosa* spp. are not terminally differentiated and ca. 20% of bacteroids recovered from nodules, together with bacteria present in infection threads, can resume growth (Marchetti *et al.*, 2011). To evaluate the specific fates of mutualists and non mutualists in plants infected by a mixed population, we monitored the fitness of total nodule bacteria over time following co-inoculation of *Mimosa pudica* seedlings with a mixture (1/1 ratio, 10⁶ total bacteria/plant) of isogenic N₂-fixing and non-fixing strains of *C. taiwanensis*. Fix⁺ and Fix⁻ strains only differed by the presence of the *nifH* gene, encoding the nitrogenase reductase subunit of the nitrogenase enzyme, and of constitutively expressed GFP or mCherry fluorescent genes. For technical reasons (see Materials and methods), nodules were only collected from 14 dpi. Importantly, each nodule was individually analyzed for bacterial fitness by plating, allowing analysis at the nodule and plant individual levels. In these experimental conditions 97% of the nodules were infected by either Fix⁺ or Fix⁻ bacteria.

We observed a marked difference in the reproductive fitness of Fix⁺ and Fix⁻ bacteria from the same plant over time, which significantly differed from 21 days post-infection (dpi) and up to 28 fold on average (Figure 1A and Figure 1—figure supplement 1), perhaps because of plant control mechanisms, including sanctions (Kiers *et al.*, 2003) and possibly PFF. A significant difference was also obtained from 28 dpi when analyzing control plants singly-infected with either Fix⁺ or Fix⁻ strains (Figure 2A). Non-fixers did not proliferate better than fixers even at 14 dpi (Figure 1A) possibly because the metabolic cost paid by bacteria to fix nitrogen in terms of ATP and reducing power is too low to be detected in our experimental conditions, or because plant sanctions/PFF and the metabolic cost of nitrogen fixation equilibrate until sanctions become prominent. The resulting net fitness cost of cooperation, which is the weighted metabolic cost of nitrogen fixation by any form of plant control, thus appeared to be zero or negative, enabling mutualism to spread.

The differential fitness was not due to a better nodulation competitiveness of Fix⁺ bacteria. The number of nodules formed by each strain was indeed proportional to the inoculum ratio (1/1) throughout the time course (Figure 3), confirming that bacterial nitrogen-fixing ability is not selected at the root entry level (Hahn and Studer, 1986; Westhoek *et al.*, 2017). Yet the number of nodules in nitrogen-starved non-fixing plants (infected with 99% or 100% Fix⁻) constantly increased over a 42 day period, while this number reached a plateau at ca. 20 dpi in healthy N₂-fixing plants (infected with 50% or 100% Fix⁺) (Figure 4), indicative of a mechanism of autoregulation of nodulation acting at the whole-plant level (Ferguson *et al.*, 2010) and depending on the nitrogen status of the plant

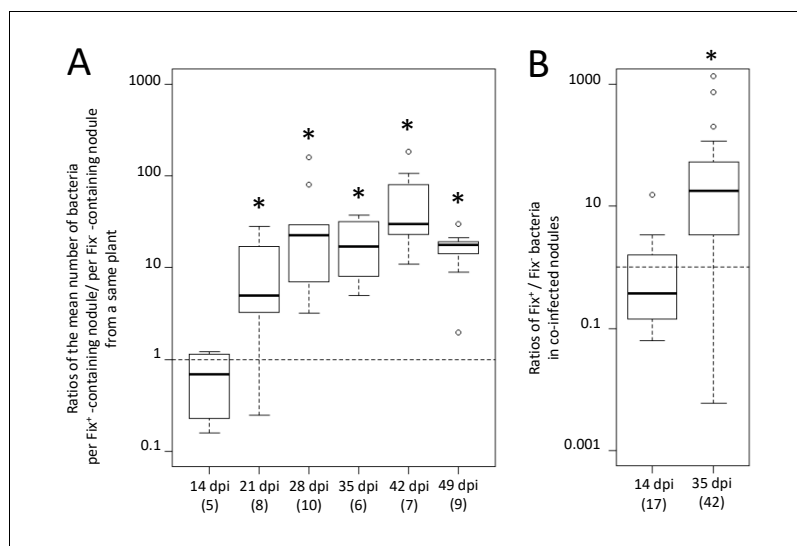


Figure 1. Kinetics of reproductive fitness of Fix^+ and Fix^- bacteria in nodules following co-inoculation of *M. pudica*. *M. pudica* plants were co-inoculated with a mixture of Fix^+ and Fix^- strains at a 1/1 ratio, using 10^6 (A) or 10^{10} bacteria/plant (B). Nodules were individually analyzed by plating their bacterial population (see **Figure 1—figure supplement 1**). Co-infected nodules represented ca. 3% (A) or 20% (B) of the nodules. (A) The ratio of the mean number of bacteria per Fix^+ -containing nodule to the mean number of bacteria per Fix^- -containing nodule was calculated for each individual plant at each time point (see **Figure 1—figure supplement 1**) and box plots represent the distribution of these ratios (**Figure 1—source data 1**). Only single-infected nodules were taken into account in this graph. (B) Box plots represent the distribution of the ratios of Fix^+ bacteria to Fix^- bacteria in co-infected nodules (**Figure 1—source data 2**). Central rectangles span the first quartile to the third quartile (that is, the interquartile range or IQR), bold segments inside rectangles show the median, unfilled circles indicate suspected outliers, whiskers above and below the box show either the locations of the minimum and maximum in the absence of suspected outlying data or $1.5 \times \text{IQR}$ if an outlier is present. Horizontal dashed lines correspond to ratios equal to 1. The number of plants (A) or nodules (B) analyzed is indicated in brackets. *Significant differences between the number of Fix^+ and Fix^- bacteria per nodule ($p < 0.05$, multiple comparison test after Kruskal-Wallis (A); $p < 0.001$, after Student t-test with paired data (B)).

DOI: <https://doi.org/10.7554/eLife.28683.003>

The following source data and figure supplement are available for figure 1:

Source data 1. Reproductive fitness of nodule bacteria following co-inoculation with Fix^+ (CBM2700) and Fix^- (CBM2707) *C. taiwanensis*.

DOI: <https://doi.org/10.7554/eLife.28683.005>

Source data 2. Reproductive fitness of nodule bacteria in nodules co-infected by Fix^+ (CBM2700) and Fix^- (CBM2707) *C. taiwanensis*.

DOI: <https://doi.org/10.7554/eLife.28683.006>

Figure supplement 1. Kinetics of reproductive fitness of nodule bacteria following co-inoculation with Fix^+ (CBM2700) and Fix^- (CBM2707) *C. taiwanensis*.

DOI: <https://doi.org/10.7554/eLife.28683.004>

(Malik et al., 1987; van Noorden et al., 2016). This difference in time course increases the chance that a rare Fix^+ among a Fix^- population will form a nodule.

To identify the spatial level at which selection applies we first analyzed double occupancy nodules, which were obtained in significant proportion by modifying the plant culture system and increasing the inoculum density by four logs (see Materials and methods). Co-infected nodules contained a similar number of Fix^+ and Fix^- bacteria at 14 dpi, but on average ca. 80 times more N_2 -fixing bacteria than non-fixing bacteria at 35 dpi (**Figure 1B**), indicating that the control occurs at the nodule scale. Previous studies established that bacteroids do not persist in nodule cells of nitrogen-starved plants infected only by non-fixers, leading to premature nodule senescence (Berrabah et al., 2015; Hirsch and Smith, 1987), while they persist in healthy plants singly-infected with fixers. We therefore then analyzed the viability of bacteroids on sections of singly-occupied or double-occupied nodules collected from co-inoculation experiments using propidium iodide (PI), which stains dead

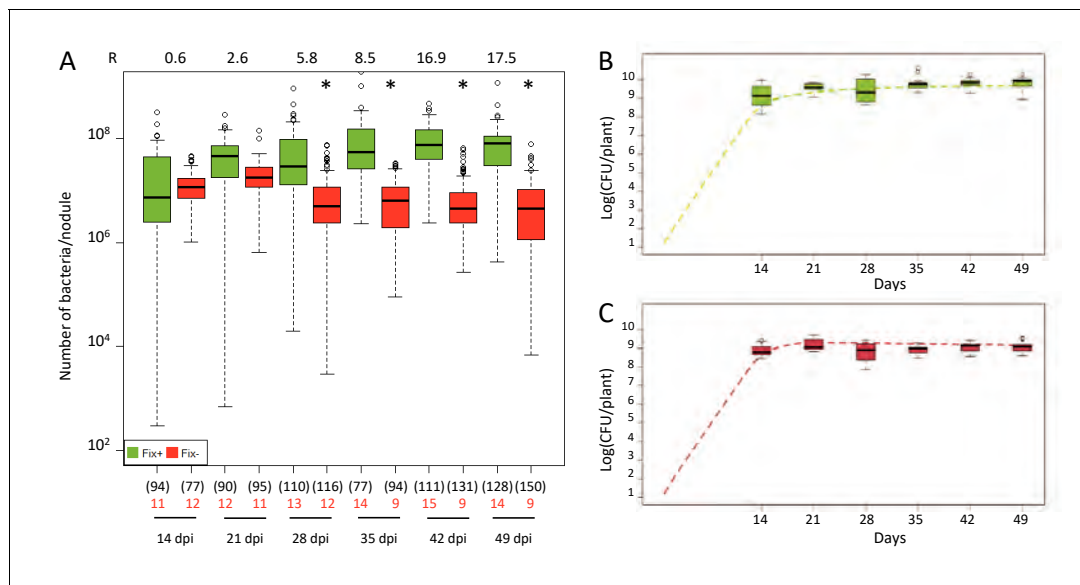


Figure 2. Kinetics of reproductive fitness of Fix^+ or Fix^- nodule bacteria following single-inoculation of *M. pudica*. (A) Fix^+ (CBM382) or Fix^- (CBM2568) *C. taiwanensis* were inoculated on *M. pudica*. Box plots represent the distribution of the number of bacteria recovered per nodule on plates. Box plots were constructed as described in **Figure 1**. R, ratios of the median number of Fix^+ bacteria per nodule on the median number of Fix^- bacteria per nodule. The number of nodules analyzed at each time point is indicated in brackets. The number of plants analyzed at each time point is indicated in red. Results are from two independent experiments (**Figure 2—source data 1**). *Significantly different from the number of Fix^+ bacteria per nodule ($p < 0.05$ multiple comparison test after Kruskal-Wallis). (B, C) Theoretical reproductive fitness of Fix^+ (B) and Fix^- bacteria (C) following single-inoculation of *M. pudica* as compared to experimental data. Dotted lines represent bacterial populations per plant averaged over 200 replicate simulations (**Figure 2—source data 2**). Box plots represent the distribution of the number of bacteria experimentally recovered per plant. Experimental data are from (A).

DOI: <https://doi.org/10.7554/eLife.28683.007>

The following source data is available for figure 2:

Source data 1. Reproductive fitness of nodule bacteria following single-inoculations with either Fix^+ (CBM382) or Fix^- (CBM2568) *C. taiwanensis*.

DOI: <https://doi.org/10.7554/eLife.28683.008>

Source data 2. Simulation data for the reproductive fitness of Fix^+ and Fix^- bacteria following single inoculations of *M. pudica*.

DOI: <https://doi.org/10.7554/eLife.28683.009>

cells (Virta et al., 1998). Bacteroid viability in Fix^+ -occupied nodules remained stable from 14 to 35 dpi (**Figure 5D**). By contrast, bacteroids in the nitrogen-fixing zone of Fix^- -occupied nodules started losing viability at 16–21 dpi and were all dead (PI-stained) at 35 dpi (**Figure 5E**). Electron microscopy confirmed signs of nodule cell and bacterial degeneration in Fix^- -occupied nodules at 19 dpi (**Figure 6**). Co-infected nodules showed clear sectoring, with infected plant cells in one part filled with Fix^+ strains and in the other part filled with Fix^- strains (**Figure 5FGHI**). We never observed co-infected nodule cells. While at 14 dpi both strains were alive (**Figure 5G**), at 35 dpi only Fix^- bacteroids were PI-stained confirming that Fix^+ and Fix^- intracellular bacteria have distinct fates within the same nodule (**Figure 5HI**). The ca. 5×10^6 bacteria recovered at 35 dpi from nodules infected with only Fix^- bacteria may thus be bacteria colonizing the infection threads and the infection zone.

In conclusion we provide evidence for differential spatio-temporal dynamics of N_2 -fixing and non-fixing partners during the symbiotic process, highlighting the importance of considering temporal variations when studying the evolution of cooperative interactions (Barker and Bronstein, 2016). We established that the control of mutualism (i) acts at the nodule cell scale, (ii) occurs relatively early, ca. 16–21 days after inoculation when the wild-type nitrogenase is fully active in Fix^+ bacteria (**Figure 7**) and (iii) leads to up a ca. 80 fold relative increase in mutualistic partners.

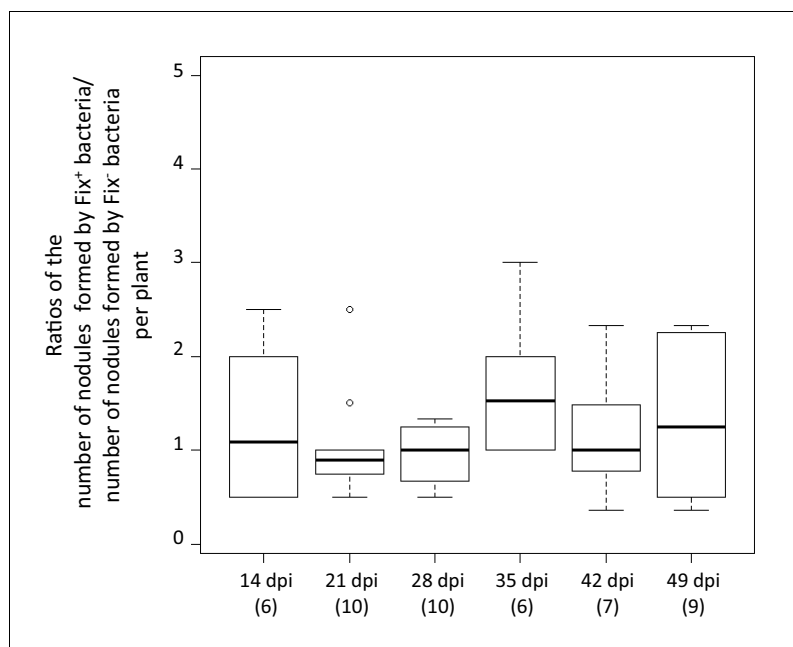


Figure 3. Relative number of nodules formed by Fix^+ and Fix^- bacteria per plant individual. *M. pudica* plants were co-inoculated with the CBM2700 (Fix^+ , GFP) and CBM2707 (Fix^- , mCherry) strains at a 1/1 ratio. The number of plants analyzed for each time point is indicated in brackets. Boxplots were constructed as described in **Figure 1**. No significant differences were observed between the number of nodules formed by Fix^+ bacteria and Fix^- bacteria per plant at the different time points ($p > 0.05$, Student *t*-test with paired data at each time point or multiple comparison test after Kruskal-Wallis on the whole dataset) (**Figure 3—source data 1**).

DOI: <https://doi.org/10.7554/eLife.28683.010>

The following source data is available for figure 3:

Source data 1. Relative number of nodules formed by Fix^+ and Fix^- bacteria per plant individual.

DOI: <https://doi.org/10.7554/eLife.28683.011>

Eco-evolutionary dynamics of N_2 -fixers and non-fixers through serial nodulation cycles

Next, we addressed the question of whether mutualism control will allow a minority Fix^+ subpopulation to invade the symbiotic population.

We first used our experimental data to develop a stochastic mathematical model qualitatively simulating the fate of *C. taiwanensis* populations during nodulation in *M. pudica* plants. The two key components of this model are (i) the kinetics of nodule formation from bacteria randomly chosen from the rhizospheric population and (ii) bacterial multiplication within nodules, according to bacterial genotype (see Materials and methods and **Table 1** for details on model construction and parameterization). While the model is developed as a proof-of-concept, instead of a simple deterministic model we chose to include stochasticity in the nodulation process in order to reflect the variability observed in the experimental data. In order to test our model, we first simulated the reproductive fitness of nodule bacteria following single-inoculation with either Fix^- or Fix^+ bacteria over a 49 day-period, and compared this simulation to the kinetics experimentally observed (**Figure 2BC**). We then both simulated and experimentally determined the relative proportion of Fix^+ bacteria recovered from plants co-inoculated with a minor subpopulation of Fix^+ (1%) and a major subpopulation of Fix^- (99%) bacteria over 49 days (**Figure 8**). Simulation outcomes qualitatively matched the dynamics of bacterial populations observed experimentally (**Figures 2BC** and **8**), indicating that the experimentally measured and inferred model parameters are appropriate for studying the evolutionary dynamics of *C. taiwanensis* populations in different ecological conditions.

We then used this model to explore how plant population size and the length of inoculation cycles impact on the dynamics of *C. taiwanensis* populations during serial cycles of inoculation of *M. pudica* plants and re-isolation of bacteria from nodules. Starting with a fixed proportion of Fix^+

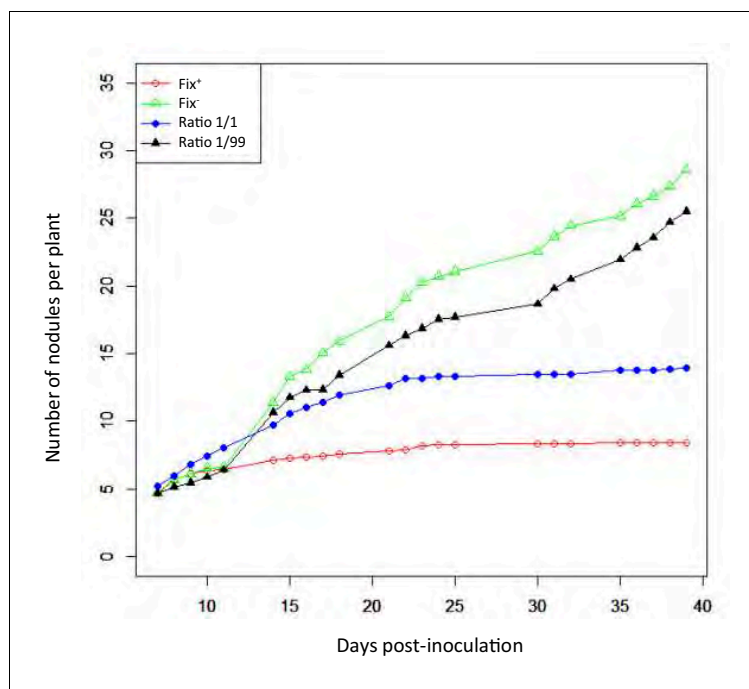


Figure 4. Nodulation kinetics. *M. pudica* plants were single-inoculated with either CBM832 (Fix⁺) or CBM2568 (Fix⁻) or co-inoculated with a mixture of both strains at a 1/1 or 1/99 ratio. First nodules appeared at 5–7 dpi (Figure 4—source data 1).

DOI: <https://doi.org/10.7554/eLife.28683.012>

The following source data is available for figure 4:

Source data 1. Nodulation kinetics of Fix⁺ (CBM382) and Fix⁻ (CBM2568) *C. taiwanensis* following single- or co-inoculation of *M. pudica*.

DOI: <https://doi.org/10.7554/eLife.28683.013>

bacteria (1% or 0.1%) in the inoculum, we varied the number of inoculated plants from 1 to 100 (or 1 to 1000) and the length of nodulation cycles (time from plant inoculation to nodule bacteria harvesting) from 14 to 49 days, which is shorter than the lifespan of a nodule in nature. We found that larger plant pools and longer cycles progressively reduced extinction probabilities and increased the proportion of Fix⁺ in the nodule bacterial population (Figure 9A and Figure 9—figure supplement 1). For example, the model predicted that using an initial inoculum of 1% Fix⁺, 4 cycles of 42 days with pools of 20 plants were sufficient to yield more than 85% of Fix⁺ bacteria in all replicates where Fix⁺ populations avoided extinction (89 times out of 100 replicates in Figure 9A). Smaller plant pools or shorter cycles all yielded higher probabilities of extinction and decreased proportions of Fix⁺ bacteria. An initially lower Fix⁺ proportion (0.1%) could be compensated for by a higher plant population size and/or a longer cycle length (Figure 9—figure supplement 1). We analyzed in detail the dynamics of Fix⁺ subpopulations over 10 cycles in a situation where the cycle length had a major impact on the evolutionary outcome (20 plants) (Figure 9A) and plotted the proportion of Fix⁺ bacteria recovered after each cycle, for cycles ranging from 14 to 49 days (Figure 9B). We observed that, in the vast majority of cases, the fate of Fix⁺ populations is already determined after the first cycle: these populations are either bound to extinction (with a probability indicated in Figure 9A) or to a gradual increase in frequency that ultimately leads to fixation. This result holds true for all cycle lengths except 14 days, where population dynamics is dominated by drift due to the equivalent fitness of Fix⁻ and Fix⁺ clones (Figure 1A). A key factor controlling the early bifurcation between extinction and fixation of Fix⁺ population is the probability that a Fix⁺ bacterium forms a nodule during the first cycle, which depends on both the size of plant pools and the length of nodulation cycles.

Understanding the influence of plant pool size is straightforward. Very few nodules are produced on each plant, creating a bottleneck in bacterial population size at each nodulation cycle. Whatever

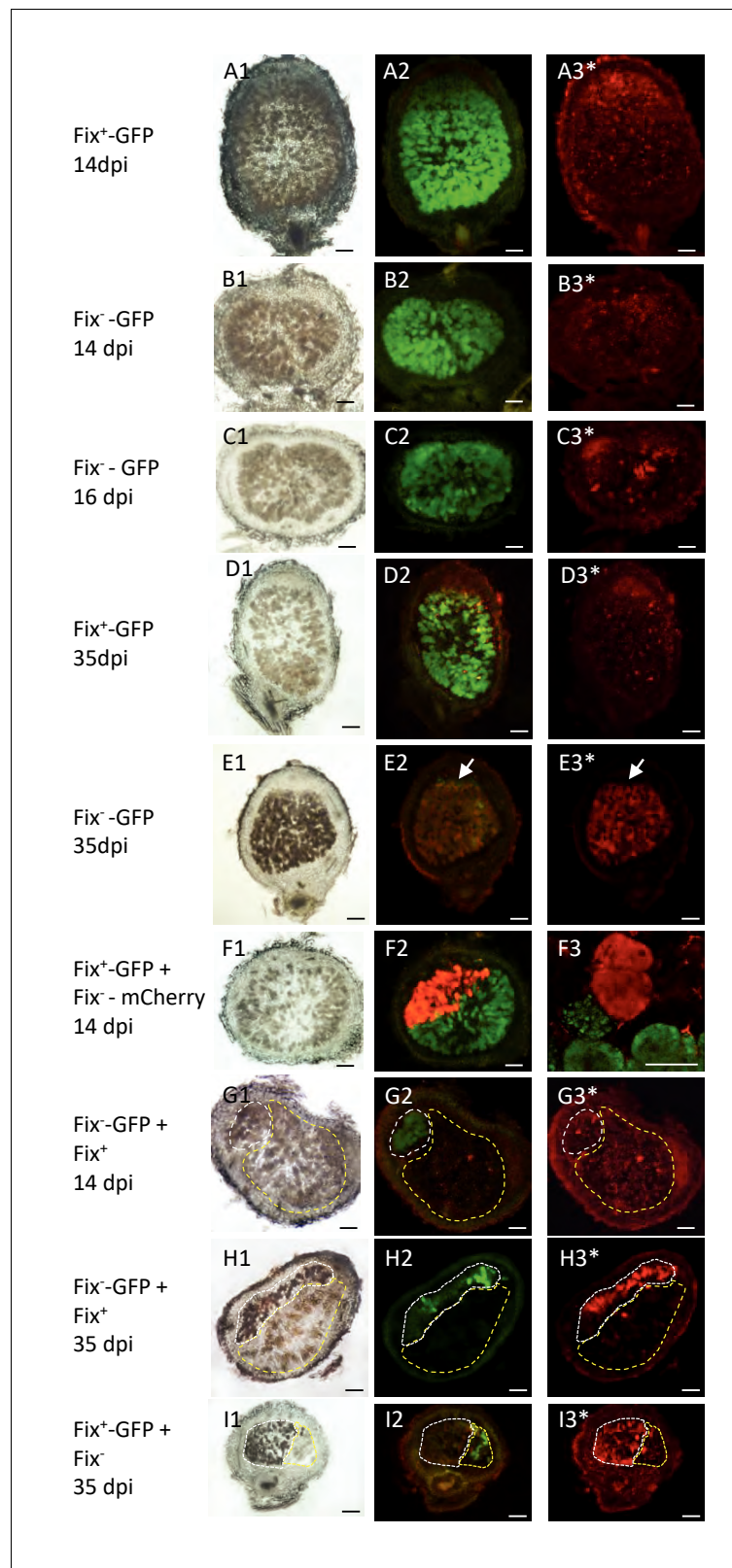


Figure 5. Viability of *Fix*⁺ and *Fix*⁻ bacteroids. *M. pudica* were co-inoculated with *Fix*⁺ and *Fix*⁻ *C. taiwanensis* at a 1/1 ratio and sections of nodules collected at 14 dpi (ABFG), 16 dpi (C) or 35 dpi (DEHI) were observed under bright field (panels 1) or fluorescent microscopy (panels 2 and 3), and after PI staining (panels with an *). Panels with the same letters represent the same nodule section. (F3), magnification of (F2) visualized by confocal
 Figure 5 continued on next page

Figure 5 continued

microscopy. (A) and (D), sections of nodules infected with a GFP-labeled Fix^+ strain. (B) (C) and (E), sections of nodule infected with a GFP-labeled Fix^- strain. (F), nodule co-infected with a GFP-labeled Fix^+ and a mCherry-labeled Fix^- strain. (G) and (H), nodules co-infected with a GFP-labeled Fix^- and an unlabeled Fix^+ strain. (I), nodules co-infected with a GFP-labeled Fix^+ and an unlabeled Fix^- strain. The white and yellow dotted lines in (GHI) delimit the areas occupied by the Fix^- and Fix^+ strains in a co-infected nodule, respectively. Note that neither the Fix^+ (D3) nor the Fix^- bacteroids (B3G3) are red-labeled by PI staining at 14 dpi whereas a few cells are PI-stained in the Fix^- -occupied nodule at 16 dpi ([C3], arrows), and Fix^- are mostly PI-labeled (dead) at 35 dpi (E3H3I3). Note that bacteria of the infection zone are still alive at 35 dpi (arrow, E2E3). Note that nodule cells filled with Fix^- are browner than nodule cells filled with Fix^+ (G1H1I1). Scale bars correspond to 100 μm except for F3 (30 μm).

DOI: <https://doi.org/10.7554/eLife.28683.014>

the cycle length, larger numbers of plants per pool increase the likelihood that at least one Fix^+ clone is sampled from the rhizospheric population, giving Fix^+ subpopulations an opportunity to increase in frequency and avoid extinction in the next cycle (Figure 9A and Figure 9—figure supplement 1). Under longer cycles, extinction probability decreases (Figure 9AB) since more nodules are produced (Figure 4) and the size of Fix^+ populations increases at a faster rate (Figure 9B) as a result

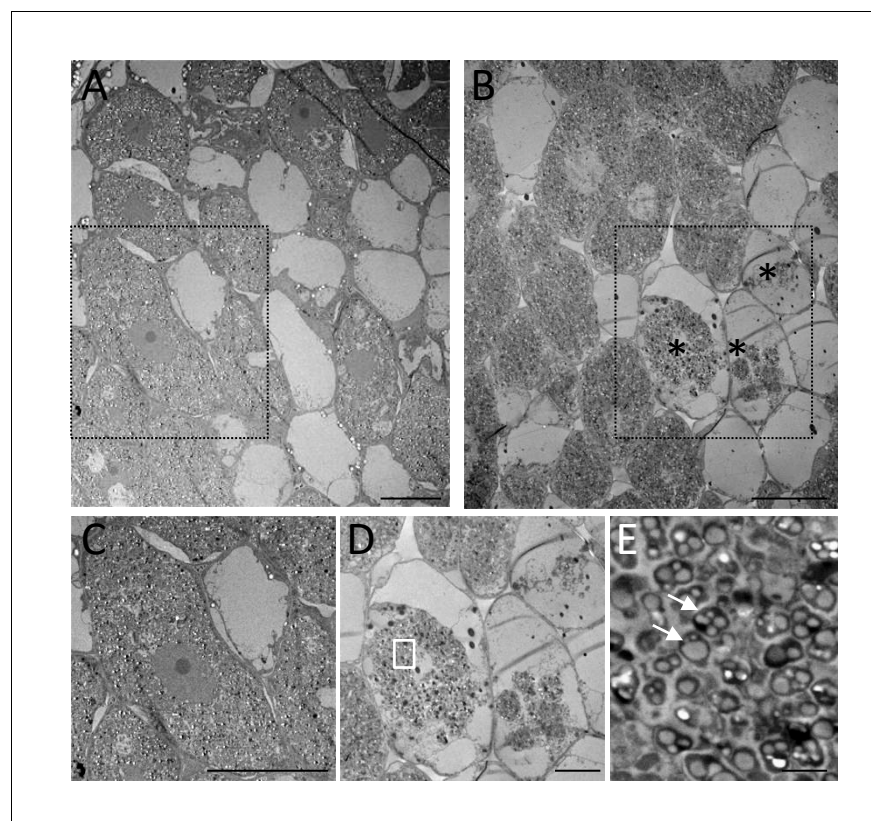


Figure 6. Electron microscopy of Fix^+ - and Fix^- -occupied nodules. *M. pudica* plants were co-inoculated with Fix^+ (CBM2708, mCherry) and Fix^- (CBM2568, unlabeled) *C. taiwanensis* at a 1/1 ratio. Nodules collected at 19 dpi (ABCDE) were sorted for mCherry expression under fluorescence microscopy and used for electron microscopy observation. Degenerated nodule cells (*) were observed in Fix^- -occupied nodules (BDE) but not in Fix^+ -occupied nodules (AC). (C) and (D) represent magnification of the zones delimited by a black dashed rectangle in (A) and (B) respectively. (E) magnification of the white rectangle in (D) showing degenerated bacteria (arrows). Scale bars represent 20 μm (ABC), 10 μm (D) and 2 μm (E).

DOI: <https://doi.org/10.7554/eLife.28683.015>

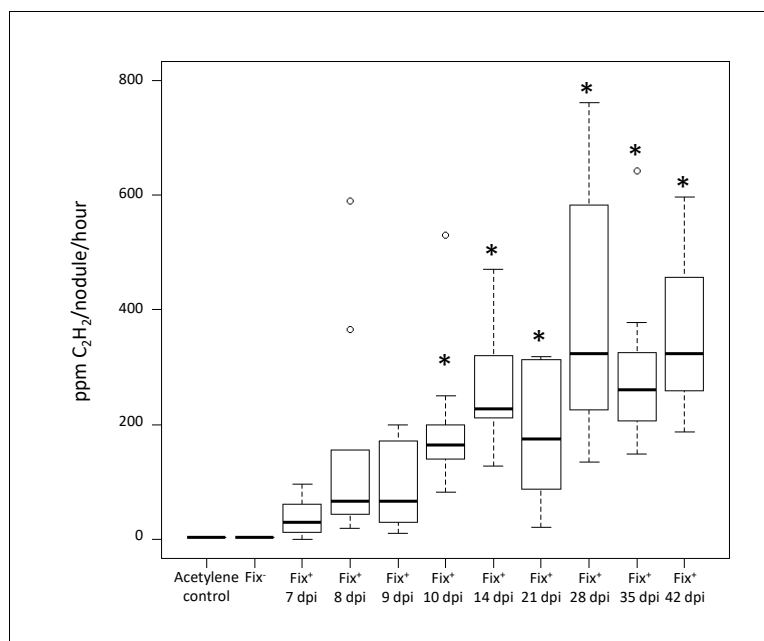


Figure 7. Kinetics of nitrogenase activity in N_2 -fixing *M. pudica* nodules. Plants were inoculated with *C. taiwanensis* CBM832 (Fix^+), and nitrogenase activity measured using the acetylene reduction assay (ARA) (**Figure 7—source data 1**). Two negative controls, i.e. tubes containing only the acetylene substrate and plants inoculated with *C. taiwanensis* CBM2568 (Fix^-), were included. In these cases, boxplots correspond to data from all time points. *, Significantly different from the negative controls ($p < 0.05$ after multiple comparison test of Kruskal-Wallis).

DOI: <https://doi.org/10.7554/eLife.28683.016>

The following source data is available for figure 7:

Source data 1. Nitrogenase activity of *C. taiwanensis* Fix^+ (CBM832).

DOI: <https://doi.org/10.7554/eLife.28683.017>

of a decrease in Fix^- fitness in older nodules (**Figure 1A**). The combined action of these two factors act on the inoculum for next cycle, generating an eco-evolutionary feedback.

To assess the predictions of the model experimentally, we performed serial inoculation-nodulation cycles of 21 or 35 days using 20 *M. pudica* plants and an initial inoculum of 5×10^3 Fix^+ / 5×10^5 Fix^- *C. taiwanensis* per plant. In each 35 day-cycle the nitrogen-fixing subpopulation increased and it reached nearly 100% of the population after four cycles (**Figure 10A**), similar to what observed with the model. Under 21 day-cycles, both simulations and experiments lead to a slower progression of Fix^+ subpopulations (**Figure 10B**). It is worth noting that an increase in frequency of the best cooperators among natural strains was also observed after three consecutive nodulation cycles between *Medicago truncatula* and *Sinorhizobium meliloti* (**Heath and Tiffin, 2009**), indicating that the selective advantage of the best N_2 -fixing strains seems to be robust to the natural diversity of symbiotic associations.

Discussion

Identifying the selective forces and ecological factors that shape mutualism is central to predicting its maintenance and dissemination over evolutionary scales. Here we provide conclusive evidence that nitrogen fixation per se, the ultimate trait that turns a parasitic rhizobium-legume association into a mutualistic one, determines the *in planta* spatio-temporal fate of endosymbiotic bacteria. Non- N_2 -fixing symbionts do not persist within cells of indeterminate nodules even when they share a nodule with N_2 -fixing symbionts, indicative of a cell autonomous senescence program as recently shown for determinate nodules (**Regus et al., 2017**). This results in the progressive and selective *in planta* expansion of fixers during the symbiotic process.

Table 1. Model parameters

Parameter	Abbreviation	Value
Size of each pool of plants*	Pool	Variable (1–1000)
Number of replicates*	Rep	Variable (5 or 100)
Length of each cycle*	Days	Variable (14–49)
Number of cycles*	Cyc	Variable (4 or 10)
Initial proportion of Fix ⁺ cells*	x	Variable (1 or 0.1)
Maximum number of new nodules/plant/day [†]	λ_{\max}	0.44
Coefficient for the auto-regulation of nodulation in nodulation kinetics [†]	a ₁	0.03
Coefficient for time-decay in nodulation kinetics [†]	a ₂	0.006
Lag for time-decay in nodulation kinetics [†]	a ₃	2
Growth rate of bacteria within nodule [†]	r	1.95
Fitness cost of nitrogen fixation [†]	c	0
Sanctions for Fix ⁺	s	1.65
Day at which additional sanctions begin [†]	ds	17
Nodule carrying capacity [†]	K	1.4×10^8

*parameters varied in the simulations

[†] experimentally measured parameters[‡]parameters inferred from experimental dataDOI: <https://doi.org/10.7554/eLife.28683.021>

The most likely explanation is that the plant exerts a post-infection control of N₂-fixation that overcomes the metabolic cost of nitrogen fixation paid by mutualistic bacteria. Sanctions could occur as defense responses and/or by decreasing nutrient supply to non-fixing bacteroids. Given that Fix⁻ and Fix⁺ bacteria are spatially segregated within nodules, the latter case could also result from the local degeneration of nodule cells, and be interpreted as an example of Partner Fidelity-Feedback mechanism occurring at the level of individual cells (*Shou, 2015*). Since control mechanisms prevent social dilemma –i.e. the possibility that one partner increases its own fitness by decreasing its investment in mutualism- and help cooperation persist (*Kiers and Denison, 2008; Frederickson, 2013; Sachs et al., 2004*), non-fixers do not threaten mutualism in our system. Yet the fate of strains able to fix intermediate levels of nitrogen fixation may be different. Monitoring the fitness of strains varying in their nitrogen fixation capacity would provide a more complete picture of mutualism control. Nevertheless, our results provide an additional example supporting the emerging idea that low quality rhizobial partners rarely benefit from low investment in mutualism (*Jones et al., 2015; Friese, 2012*). Plant sanctions resulting in bacterial fitness reduction were demonstrated in some rhizobium-legume systems by simulating N₂ deficiency via gas manipulation around nodules (*Kiers et al., 2003; Oono et al., 2011*), although not seen in other systems (*Marco et al., 2009; Ling et al., 2013*). That different plants may rely on different control mechanisms would not be surprising given the variety of mechanisms that lead to symbiosis with legumes (*Masson-Boivin et al., 2009*).

Experimental investigations can fuel a theoretical framework able to reframe general evolutionary questions in an ecological context (*Hoek et al., 2016*). Our qualitative model of the eco-evolutionary dynamics of mutualistic and non-mutualistic populations includes serial inoculation-nodulation cycles. This regime mimics an experimental set up of horizontal transmission of rhizobia across plant generations albeit on an accelerated basis. A general outcome of the model is that rare fixers will invade a population dominated by non-fixing bacteria, above a threshold combination of plant and bacterial population sizes and cycle lengths. The model helps explore further combinations of number of cycles, cycle lengths and plant pool sizes to hypothesize the evolutionary trajectory of the Fix⁺ genotype. While the selective advantage of the Fix⁺ phenotype is expected to ensure its fixation in a deterministic manner, strong population bottlenecks occurring at the nodulation step introduce a source of stochasticity in these dynamics and may thus prevent the action of directional selection. The effect of stochasticity has been shown to be of immense evolutionary consequence in related

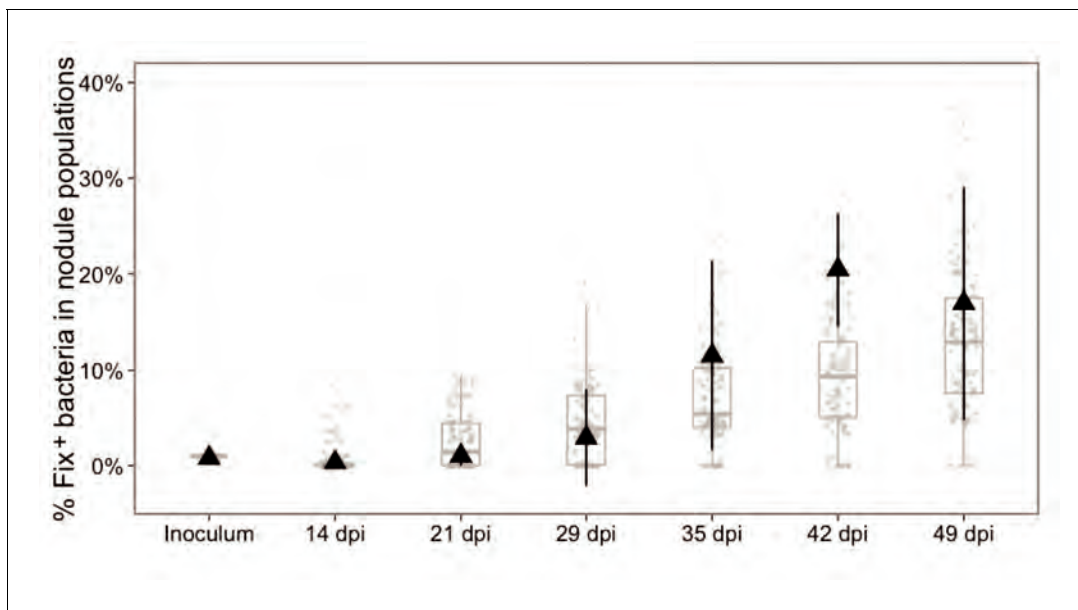


Figure 8. Experimental and theoretical reproductive fitness of Fix^+ and Fix^- bacteria following co-inoculation of *M. pudica* (ratio 1/100). The proportion of Fix^+ clones in nodules was experimentally measured and simulated over 49 days, following co-inoculation of 20 plants. Experimental data are shown as black triangles (**Figure 8—source data 1**). Black error bars represent standard deviation from 2 to 3 replicates. The results from 100 replicate simulations are shown as grey dots and boxplots (**Figure 8—source data 2**).

DOI: <https://doi.org/10.7554/eLife.28683.018>

The following source data is available for figure 8:

Source data 1. Experimental data for the reproductive fitness of Fix^+ and Fix^- bacteria following co-inoculation of *M. pudica* (ratio 1/100) over 49 days.

DOI: <https://doi.org/10.7554/eLife.28683.019>

Source data 2. Simulation data for the reproductive fitness of Fix^+ and Fix^- bacteria following co-inoculation of *M. pudica* (ratio 1/100) over 49 days.

DOI: <https://doi.org/10.7554/eLife.28683.020>

models of host parasite coevolution (*Papkou et al., 2016*). Another characteristic of our system is that, when the Fix^+ populations increase in abundance then so does their proliferation, leading to a quick increase of Fix^+ over successive nodulation cycles (**Figure 9B**). This interaction between the demographic composition of the population and the evolutionary success of one of the traits is an example of the eco-evolutionary feedback present in this system.

Although the selective and ecological forces at play in the lab and in field conditions may differ significantly, our results predict that both forces have played a major role in the evolution of the rhizobium-legume mutualism by favoring the fixation of emerging N_2 -fixing sub-populations among uncooperative symbiotic populations as well as their evolutionary maintenance. Yet the uncooperative population does not become extinct within nodules, likely because sanctions mainly target bacteroids of the nitrogen fixation zone. Releasing non-fixing bacteria may allow progenitors to meet appropriate hosts or to evolve new symbiotic traits. This loose selection process helps maintain genetically diverse rhizobial communities in the soil and shape the ecology and evolution of rhizobia. More generally, acknowledging the existence of non-cooperators as an integral component of the ecological and evolutionary dynamics of mutualistic interactions may provide a better understanding of the long-term persistence of bacterial lineages (*Heath and Tiffin, 2009; Heath and Stinchcombe, 2014; Tarnita, 2017; Fiegna et al., 2006; Hammerschmidt et al., 2014*).

An emerging trend in fundamental and applied plant microbiology is to select upon microbes indirectly through the host (*Mueller and Sachs, 2015*). This engineering approach, called host-mediated selection, involves selection of microbial traits that are not selectable in vitro. Modelling the eco-evolutionary scenarios provides predictions to guide experimental evolution studies aiming at designing beneficial microbes (*Marchetti et al., 2010; Marchetti et al., 2017*) and microbiomes (*Mueller and Sachs, 2015; Johns et al., 2016*).

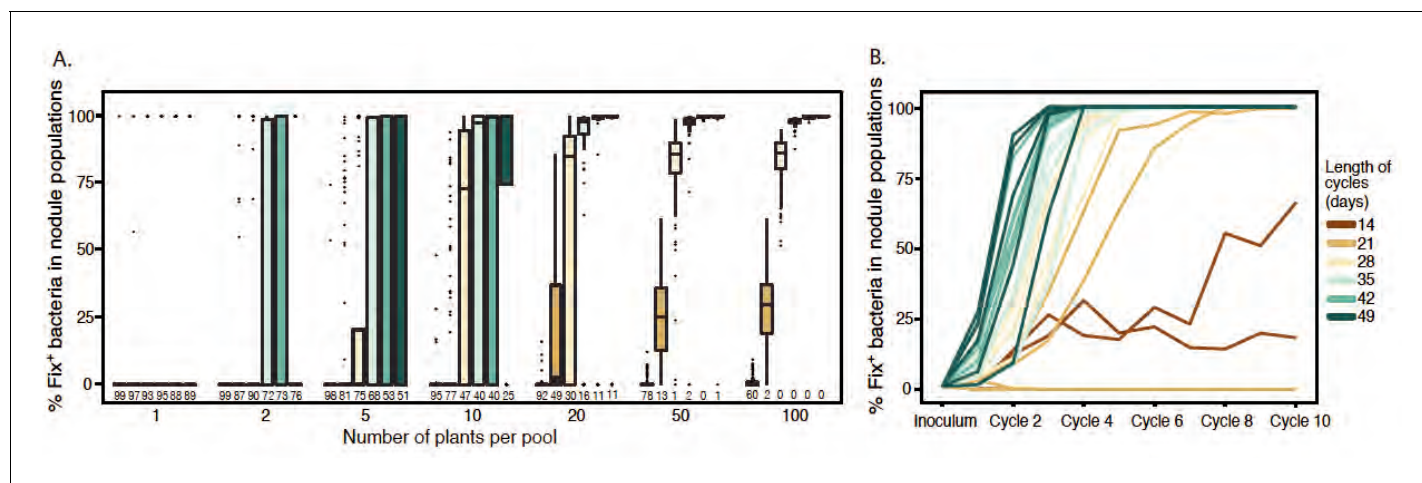


Figure 9. Effect of cycle length and plant numbers on the predicted distributions of Fix⁺ population sizes. Model simulations were performed with an initial proportion of 1% Fix⁺ in the bacterial population inoculated to a pool of plants. The length of each cycle and the number of plants per pool varied as indicated in the legend. (A) Final proportion of Fix⁺ clones after four cycles (Figure 9—source data 1). Boxplots represent the distribution of the final proportion of Fix⁺ clones from 100 simulations. The length of inoculation cycles ranged from 14 to 49 days and the number of plants per pool from 1 to 100. Numbers underneath each boxplot indicate the number of replicate simulations where Fix⁺ sub-populations became extinct after four cycles. (B) Increase in the proportion of Fix⁺ clones along 10 inoculation cycles of 14, 21, 28, 35, 42 or 49 days (Figure 9—source data 2). The number of plants per pool was 20. Representative trajectories of 5 replicate pools are shown in each case.

DOI: <https://doi.org/10.7554/eLife.28683.022>

The following source data and figure supplement are available for figure 9:

Source data 1. Simulation data for the final proportion of Fix⁺ bacteria after four inoculation cycles.

DOI: <https://doi.org/10.7554/eLife.28683.024>

Source data 2. Simulation data for the increase in proportion of Fix⁺ bacteria along 10 cycles.

DOI: <https://doi.org/10.7554/eLife.28683.025>

Source data 3. Simulation data for the effect of cycle length and plant number on the Fix⁺ population sizes after four cycles.

DOI: <https://doi.org/10.7554/eLife.28683.026>

Figure supplement 1. Effect of cycle length and plant numbers on the predicted distribution of Fix⁺ population sizes.

DOI: <https://doi.org/10.7554/eLife.28683.023>

Materials and methods

Bacterial strains and growth conditions

Strains and plasmids used in this study are listed in **Table 2**.

C. taiwanensis strains were grown at 28°C on TY medium supplemented with 6 mM CaCl₂ and 200 µg/ml streptomycin. *E. coli* strains were grown at 37°C on LB medium and antibiotics were used at the following concentrations: kanamycin 25 µg/ml, trimethoprim 100 µg/ml, tetracycline 10 µg/ml. For in vitro competition experiments, strains were pre-cultured in TY medium, mixed in equal proportion then co-inoculated to a 100 ml culture in TY medium. Bacteria were plated every 2 hr during the exponential phase, at the entry of stationary phase and 15 hr after the entry into the stationary phase. Plated bacteria were grown for 48 hr at 28°C then green and red bacteria were counted using a fluorescence stereo zoom microscope (Axiozoom V16, Zeiss).

Mutant construction

Mutant and labeled strains of *C. taiwanensis* were constructed using the mutagenesis system developed by Flannagan et al. (Flannagan et al., 2008) involving the suicide plasmid pGPI-SceI carrying an I-SceI recognition site and the pDAI-SceI replicative plasmid expressing the I-SceI nuclease. To construct the unmarked *C. taiwanensis nifH* mutant, regions upstream and downstream *nifH* were amplified with the oCBM1821-oCBM2362 and oCBM1822-oCBM2363 primer pairs using GoTaq DNA polymerase (Promega). The two PCR products were digested with XbaI-BamHI and BamHI-EcoRI respectively and cloned into the pGPI-SceI plasmid digested by XbaI and EcoRI. Ligation

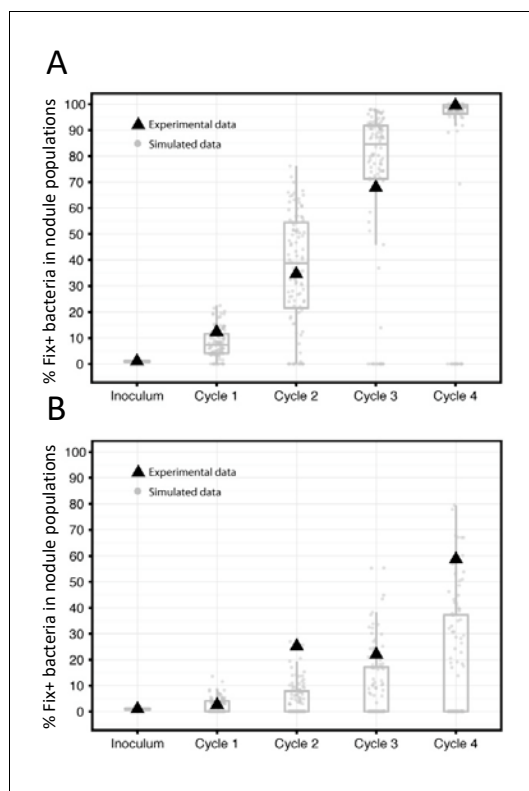


Figure 10. Frequency of Fix⁺ bacteria over 4 cycles of 35 (A) or 21 (B) days: simulations and experimental validation. The proportion of Fix⁺ clones over four inoculation cycles was simulated and measured experimentally. Simulations and experiment were performed with an initial proportion of Fix⁺ clones of 1% and pools of 20 plants. Experiments were performed with an inoculum of 5×10^3 Fix⁺/ 5×10^5 Fix⁻ *C. taiwanensis* per plant. The results from 100 replicate simulations are shown as grey dots and boxplots (**Figure 10—source data 1**). Experimental data are shown as black triangles (**Figure 10—source data 2**).

DOI: <https://doi.org/10.7554/eLife.28683.027>

The following source data is available for figure 10:

Source data 1. Simulation data for the frequency of Fix⁺ bacteria over 4 cycles of 35 or 21 days.

DOI: <https://doi.org/10.7554/eLife.28683.028>

Source data 2. Experimental data for the frequency of Fix⁺ and Fix⁻ bacteria over 4 cycles of 35 or 21 days.

DOI: <https://doi.org/10.7554/eLife.28683.029>

not noticeably affect bacterial growth rate.

Oligonucleotide sequences used for genetic constructions are provided in **Supplementary file 1**.

Plant tests

Mimosa pudica seeds were of Australian origin (B and T World Seed, Pagnignan, France) and were sterilized as described (**Chen et al., 2003**). Seedlings were cultivated in Gibson tubes (2 *M. pudica* plantlets/tube) as previously described (**Marchetti et al., 2014**). To increase the frequency of coinfection, plants were grown on 12 cm² plates (three plants per plate) containing slanting nitrogen-

products were transformed into a DH5 α λ pir *E. coli* strain. The resulting plasmid was transferred into *C. taiwanensis* CBM832 by triparental mating using pRK2013 as helper plasmid. Transconjugants that have integrated the plasmid by single crossing over were selected on streptomycin and trimethoprim and verified by PCR using the oCBM1824-oCBM2363 and oCBM1825-oCBM2362 primer pairs. Then we introduced the pDAI-Scel replicative plasmid into these strains by conjugation and selection on tetracyclin. Expression of the I-Scel nuclease causes a double strand break into the inserted plasmid and promotes DNA recombination. Mutants deleted in *nifH* were screened by trimethoprim sensitivity and verified by PCR using the oCBM1824-oCBM1825 pair of primers. Mutants were then cultivated on unselective TY medium. Tetracycline sensitive colonies which have lost the pDAI-Scel plasmid were selected.

The Pps-GFP and Pps-mCherry fusions were inserted into the wild-type and *nifH* mutant of *C. taiwanensis* at the same chromosomal locus, i. e. in the intergenic region between the *glmS* and RALTA_A0206 genes using the same pGPI-Scel/pDAI-Scel mutagenesis system. Flanking regions of the insertion site were amplified by PCR using the Phusion DNA polymerase (ThermoFisher Scientific) and the oCBM2619-oCBM2620 and oCBM2621-oCBM2622 primer pairs. PCR products were digested by XbaI and Acc65I or Acc65I and EcoRI respectively and cloned into the pGPI-Scel plasmid digested by XbaI and EcoRI. The two fusions Pps-GFP and Pps-mCherry were obtained by digesting the pRCK-Pps-GFP and pRCK-Pps-mCherry by AvrII and SpeI and cloned into the pGPI-Scel carrying the intergenic region *glmS*-RALTA_A0206 digested by the same enzymes. The resulting pCBM161 and pCBM162 were first transformed into a DH5 α λ pir *E. coli* strain then transferred into *C. taiwanensis* by triparental mating with the pRK2013 helper plasmid. Integration of the fusions by double crossing over was carried out using the pDAI-Scel plasmid as described above. CBM2700 (Fix⁺, GFP) and CBM2707 (Fix⁻, mCherry) had the same plating efficiency in in vitro competition experiments, indicating that these genetic modifications did

Table 2. Strains and plasmids used in this study

Bacterium	Strain	Relevant characteristics	Reference/source
<i>E. coli</i>	DH5 α	<i>F recA lacZM15</i>	Bethesda research laboratory
	DH5 α λ pir	<i>F recA lacZM15</i> λ pir	HP Schweizer
<i>C. taiwanensis</i>	LMG19424	Wild-type strain isolated from <i>Mimosa pudica</i> in Taiwan	(Chen et al., 2001)
	CBM832	LMG19424 derivative resistant to Streptomycin, Str ^R	M. Hynes
	CBM2568	CBM832 deleted in <i>nifH</i> , Str ^R	This study
	CBM2700	CBM832 carrying a Pps-GFP fusion downstream <i>glmS</i> , Str ^R	This study
	CBM2701	CBM2568 carrying a Pps-GFP fusion downstream <i>glmS</i> , Str ^R	This study
	CBM2707	CBM2568 carrying a Pps-mCherry fusion downstream <i>glmS</i> , Str ^R	This study
	CBM2708	CBM832 carrying a Pps-mCherry fusion downstream <i>glmS</i> , Str ^R	This study
Plasmids	Name	Relevant characteristics	Reference/source
	pGPI-Scel	<i>ori</i> _{R6K} , <i>mob</i> ⁺ , carries a I-Scel site, Tri ^R	(Flannagan et al., 2008)
	pDAI-Scel	<i>ori</i> _{pBBR1} , <i>mob</i> ⁺ , carries the I-Scel gene, Tet ^R	(Flannagan et al., 2008)
	pRCK-Pps-GFP	Plasmid carrying the <i>psbA</i> promoter region fused to GFP, Kan ^R	M. Valls
	pRCK-Pps-mCherry	Plasmid carrying the <i>psbA</i> promoter region fused to mCherry, Kan ^R	M. Valls
	pCBM156	pGPI-Scel carrying the <i>nifH</i> 5' and 3' regions, Tri ^R	This study
	pCBM161	pGPI-Scel carrying the <i>glmS</i> -Ralta_A0206 intergenic region interrupted by a Pps-GFP fusion, Tri ^R	This study
	pCBM162	pGPI-Scel carrying the <i>glmS</i> -Ralta_A0206 intergenic region interrupted by a Pps-mCherry fusion, Tri ^R	This study
	pRK2013	Helper plasmid, Kan ^R	(Figurski and Helinski, 1979)

Str, streptomycin; Tri, trimethoprim; Tet, tetracycline; Kan, kanamycin.

DOI: <https://doi.org/10.7554/eLife.28683.030>

free Fahraeus agar medium for 3 days at 28°C. Roots were covered with a sterile, gas-permeable, and transparent plastic film (BioFolie 25; Sartorius AG, Vivascience, Bedminster, NJ, U.S.A.). For single-strain inoculation experiments, each plant in Gibson tubes was inoculated with 5.10^5 bacteria either CBM832 (wild-type) or its isogenic *nifH* mutant, CBM2568. For co-inoculation experiments in Gibson tubes, plants were inoculated with the two isogenic strains CBM2700 (wild-type, GFP labeled) and CBM2707 (*nifH*, mCherry labeled) at ratio 1/1 (5.10^5 bacteria of each strain per plant) or 1/100 (5.10^3 bacteria of CBM2700 and 5.10^5 bacteria of CBM2707 per plant). For co-inoculation experiments in plates, plants were inoculated with 10^{10} bacteria of each strain per plant.

To measure the number of nodule bacteria over time, all nodules from 5 to 10 individual plants, except very small nodules, were individually collected with at least 2 mm of root left on both sides of nodules and treated at each time point. We did not collect very small nodules since there was a risk that the sterilization agents penetrate these nodules. In the same line we did not collect nodules before 14 dpi since most nodules were very small at that stage. Nodules were surface sterilized for 15 min in a 2.5% sodium hypochlorite solution, rinsed with water and crushed. Each nodule crush was diluted and plated using an easy spiral automatic plater (Interscience). Colonies were counted after 2 day-incubation at 28°C, under a fluorescence stereo zoom microscope (Axiozoom V16, Zeiss) when appropriate.

For nodulation kinetics, the number of nodules formed on 20 plants grown in Gibson tubes was counted daily for 6 weeks.

For serial inoculation-nodulation cycles on *M. pudica* plants, 10 Gibson tubes of plants were inoculated with CBM2700 and CBM2707 in 1/100 ratio as described above. 35 days after inoculation, all nodules were collected, surface-sterilized and crushed together. The nodule crush was used to inoculate a new set of 10 tubes of plants with 50 μ l of a 1/10 dilution of the nodule crush per plant. At

each cycle, dilutions of the nodule crush were spread on plates, incubated 2 days at 28°C and colonies were counted under a fluorescence stereo zoom microscope.

Cytological analyses

The viability of nodule bacteria was estimated using propidium iodide staining at a concentration of 20 mM in DMSO (Molecular Probes, Fisher scientific, Oregon) on 55/58 μm nodule sections. For each experiment, a dozen nodules were individually analyzed at 14, 16, 17, 21, 28 and 35 dpi. For electron microscopy analysis, nodules were fixed in glutaraldehyde (2.5% in phosphate buffer 0.1 M [pH 7.4]), osmium treated, dehydrated in an alcohol series, and embedded in Epon 812. Semithin nodule sections were observed by brightfield microscopy after staining in 0.1% aqueous toluidine blue solution and observed under a Zeiss Axiophot light microscope. Ultrathin sections were stained with uranyl acetate and observed with a TEM Hitachi HT7700.

Acetylene reduction assays

M. pudica plants were inoculated with the wild-type strain of *C. taiwanensis* CBM832. At different time points, plants were removed from the culture Gibson tube and placed in an airtight tube and incubated with 1 ml of acetylene for 4 hr. 100 μl of gas were then injected into a gas chromatograph (Agilent GC7820). The area of the ethylene peak was measured and compared to an ethylene standard of known concentration. Ethylene background was estimated by analyzing empty tubes incubated with the same amount of acetylene.

Mathematical model and simulations

The model aimed at simulating nodulation dynamics during single or repeated inoculation-nodulation cycles. First we parameterized the population dynamics during the symbiosis process. Then we simulated repeated nodulation cycles varying the following parameters: (i) the Fix⁺/Fix⁻ ratio in the initial inoculum, (ii) the number of inoculated plants, and (iii) the cycle length. The model ran on a pool of plants (of given, variable size) from which nodules were collected and mixed together after each inoculation cycle. For each time-step (1 day) after inoculation, the number of new nodules formed on each plant was randomly drawn from a Poisson distribution of parameter $\lambda(t, \text{nod}_t^+)$, which is itself a function of time t and of the number of nodules already present on the plant nod_t^+ at time t . The maximal number of nodules that could potentially be formed per day per plant was set to λ_{max} . Changing the value of parameter λ depending on the number of Fix⁺ nodules already present on the plant simulated the autoregulation of nodulation process; this was done by subtracting the factor $a_1 \times \text{nod}_t^+$ from λ_{max} . Lastly, to allow for some 'aging' process that would decrease the rate of nodulation with time (even for plants inoculated only with Fix⁻ bacteria), we incorporated a time-decay coefficient: $a_2 \times (t - a_3)$, meaning that a reduction in the rate of nodulation occurred at a rate a_2 when $t > a_3$. This time-decay factor was set to 0 when $t < a_3$. Therefore, the parameter of the Poisson distribution controlling the rate at which new nodules are formed was given by: $\lambda(t, \text{nod}_t^+) = \lambda_{\text{max}} - a_1 \times \text{nod}_t^+$ for $t < a_3$ and by: $\lambda(t, \text{nod}_t^+) = \lambda_{\text{max}} - a_1 \times \text{nod}_t^+ - a_2 \times (t - a_3)$ for $t > a_3$. Since nodules are persistent once formed, we further set: $\lambda(t, \text{nod}_t^+) \geq 0$. Experimental evidence indicated that the number of inoculated bacteria did not affect nodulation kinetics as long as the total inoculum remains above 10^3 bacteria per plant. These conditions were met in all experiments described in this work. Therefore, we did not explicitly take inoculum size into consideration in the simulations, and restricted the applicability of our model to cases where inoculum was above this threshold value.

The second module of the model dealt with bacterial multiplication within plant nodules. Within each nodule we assumed a logistic growth model for the bacteria given by: $X(t + 1) = (r - c - \text{su}_{\text{ds}}(t)) \times X(t) \times (1 - X(t)/K)$, where r was the growth rate, c the net fitness cost of nitrogen fixation in Fix⁺ bacteria, $\text{su}_{\text{ds}}(t)$ the additional plant sanctions against Fix⁻ bacteria occurring in the later phase of the interaction, $X(t)$ the bacterial population at time t and K the nodule carrying capacity. In our simulations, we set $c = 0$ since we experimentally did not detect any difference in the populations of Fix⁻ or Fix⁺ nodule bacteria at 14 dpi. We emphasize that a net fitness cost of 0 does not necessarily imply that nitrogen fixation does not impose a metabolic burden on the bacteria (referred to as 'metabolic cost' in the results section). Instead, this burden, if significant during the early steps of the interaction, may be compensated for by plant control mechanisms acting at a basal level. Beyond this time

point, additional plant sanctions (possibly including partner fidelity-feedback) were given by $su_{ds}(t)$, taking the value s of plant sanctions indicated in **Table 1** as long as the age of the nodule was higher than ds days (denoted by the step function $u_{ds}(t)=0$ if $t < ds$ or $u_{ds}(t)=1$ if $t > ds$).

Parameters values were estimated by computing the minimal root mean square error (RMSE) of experimental data (nodulation kinetics and bacterial multiplication within nodules) versus model outputs calculated for a range of parameter values. Parameter values selected to minimize RMSE are indicated in **Table 1**. Simulations were implemented in R (*R Core Team, 2014*) and code is available in the **Source code file 1**.

Acknowledgements

We are grateful to Jacques Batut, Peter Young and Erik Hom for helpful comments on the manuscript. This work was supported by funds from the French National Research Agency (ANR-12-ADAP-0014-01 and ANR-16-CE20-0011-01), and the French Laboratory of Excellence project 'TULIP' (ANR-10-LABX-41; ANR-11-IDEX-0002-02). BD was supported by an INRA-Région Occitanie fellowship. CSG acknowledges support from the Max Planck Society.

Additional information

Funding

Funder	Grant reference number	Author
Agence Nationale de la Recherche	ANR-12-ADAP-0014-01	Marta Marchetti Catherine Masson-Boivin Delphine Capela
Institut National de la Recherche Agronomique		Benoit Daubech
Max-Planck-Gesellschaft		Chaitanya S Gokhale
Agence Nationale de la Recherche	ANR-16-CE20-0011-01	Marta Marchetti Catherine Masson-Boivin Delphine Capela
Agence Nationale de la Recherche	ANR-10-LABX-41	Benoit Daubech Marta Marchetti Cécile Pouzet Marie-Christine Auriac Catherine Masson-Boivin Delphine Capela
Agence Nationale de la Recherche	ANR-11-IDEX-0002-02	Benoit Daubech Marta Marchetti Cécile Pouzet Marie-Christine Auriac Catherine Masson-Boivin Delphine Capela

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Benoit Daubech, Marta Marchetti, Cécile Pouzet, Marie-Christine Auriac, Investigation, Methodology; Philippe Remigi, Conceptualization, Formal analysis, Investigation, Methodology, Writing—original draft, Writing—review and editing; Ginaini Doin de Moura, Investigation; Chaitanya S Gokhale, Conceptualization, Supervision, Methodology, Writing—review and editing; Catherine Masson-Boivin, Conceptualization, Supervision, Funding acquisition, Methodology, Writing—original draft, Writing—review and editing; Delphine Capela, Conceptualization, Data curation, Formal analysis, Supervision, Investigation, Methodology, Writing—original draft

Author ORCIDsPhilippe Remigi  <http://orcid.org/0000-0001-9023-3788>Chaitanya S Gokhale  <http://orcid.org/0000-0002-5749-3665>Catherine Masson-Boivin  <http://orcid.org/0000-0002-3506-3808>**Decision letter and Author response**Decision letter <https://doi.org/10.7554/eLife.28683.034>Author response <https://doi.org/10.7554/eLife.28683.035>

Additional files**Supplementary files**

- Supplementary file 1. Primers used in this study.

DOI: <https://doi.org/10.7554/eLife.28683.031>

- Source code file 1. R code used for simulations.

DOI: <https://doi.org/10.7554/eLife.28683.032>

- Transparent reporting form

DOI: <https://doi.org/10.7554/eLife.28683.033>

References

- Akçay E. 2015. Evolutionary models of mutualism. In: Bronstein J. L (Ed). *Mutualism*. New York: Oxford University Press. p. 57–76. DOI: <https://doi.org/10.1093/acprof:oso/9780199675654.003.0004>
- Archetti M, Scheuring I, Hoffman M, Frederickson ME, Pierce NE, Yu DW. 2011. Economic game theory for mutualism and cooperation. *Ecology Letters* **14**:1300–1312. DOI: <https://doi.org/10.1111/j.1461-0248.2011.01697.x>, PMID: 22011186
- Barker JL, Bronstein JL. 2016. Temporal structure in cooperative interactions: what does the timing of exploitation tell us about its cost? *PLoS Biology* **14**:e1002371. DOI: <https://doi.org/10.1371/journal.pbio.1002371>, PMID: 26841169
- Batut J, Andersson SG, O’Callaghan D. 2004. The evolution of chronic infection strategies in the alpha-proteobacteria. *Nature Reviews Microbiology* **2**:933–945. DOI: <https://doi.org/10.1038/nrmicro1044>, PMID: 15550939
- Berrabah F, Ratet P, Gourion B. 2015. Multiple steps control immunity during the intracellular accommodation of rhizobia. *Journal of Experimental Botany* **66**:1977–1985. DOI: <https://doi.org/10.1093/jxb/eru545>, PMID: 25682610
- Bever JD, Richardson SC, Lawrence BM, Holmes J, Watson M. 2009. Preferential allocation to beneficial symbiont with spatial structure maintains mycorrhizal mutualism. *Ecology Letters* **12**:13–21. DOI: <https://doi.org/10.1111/j.1461-0248.2008.01254.x>, PMID: 19019195
- Brogghammer A, Krusell L, Blaise M, Sauer J, Sullivan JT, Maolanon N, Vinther M, Lorentzen A, Madsen EB, Jensen KJ, Roepstorff P, Thirup S, Ronson CW, Thygesen MB, Stougaard J. 2012. Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding. *PNAS* **109**:13859–13864. DOI: <https://doi.org/10.1073/pnas.1205171109>, PMID: 22859506
- Chen WM, James EK, Prescott AR, Kierans M, Sprent JI. 2003. Nodulation of mimosa spp. by the beta-proteobacterium *Ralstonia taiwanensis*. *Molecular Plant-Microbe Interactions* : *MPMI* **16**:1051–1061. DOI: <https://doi.org/10.1094/MPMI.2003.16.12.1051>, PMID: 14651338
- Chen WM, Laevens S, Lee TM, Coenye T, De Vos P, Mergeay M, Vandamme P. 2001. *Ralstonia taiwanensis* sp. nov., isolated from root nodules of Mimosa species and sputum of a cystic fibrosis patient. *International Journal of Systematic and Evolutionary Microbiology* **51**:1729–1735. DOI: <https://doi.org/10.1099/00207713-51-5-1729>, PMID: 11594603
- Deakin WJ, Broughton WJ. 2009. Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nature Reviews Microbiology* **7**:312–320. DOI: <https://doi.org/10.1038/nrmicro2091>, PMID: 19270720
- Ferguson BJ, Indrasumunar A, Hayashi S, Lin MH, Lin YH, Reid DE, Gresshoff PM. 2010. Molecular analysis of legume nodule development and autoregulation. *Journal of Integrative Plant Biology* **52**:61–76. DOI: <https://doi.org/10.1111/j.1744-7909.2010.00899.x>, PMID: 20074141
- Fiegna F, Yu YT, Kadam SV, Velicer GJ. 2006. Evolution of an obligate social cheater to a superior cooperator. *Nature* **441**:310–314. DOI: <https://doi.org/10.1038/nature04677>, PMID: 16710413
- Figurski DH, Helinski DR. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *PNAS* **76**:1648–1652. DOI: <https://doi.org/10.1073/pnas.76.4.1648>, PMID: 377280

- Flannagan RS**, Linn T, Valvano MA. 2008. A system for the construction of targeted unmarked gene deletions in the genus *Burkholderia*. *Environmental Microbiology* **10**:1652–1660. DOI: <https://doi.org/10.1111/j.1462-2920.2008.01576.x>, PMID: 18341581
- Frederickson ME**. 2013. Rethinking mutualism stability: cheaters and the evolution of sanctions. *The Quarterly Review of Biology* **88**:269–295. DOI: <https://doi.org/10.1086/673757>, PMID: 24552098
- Friesen ML**. 2012. Widespread fitness alignment in the legume-rhizobium symbiosis. *New Phytologist* **194**:1096–1111. DOI: <https://doi.org/10.1111/j.1469-8137.2012.04099.x>, PMID: 22404688
- Gehlot HS**, Tak N, Kaushik M, Mitra S, Chen WM, Poweleit N, Panwar D, Poonar N, Parihar R, Tak A, Sankhla IS, Ojha A, Rao SR, Simon MF, Reis Junior FB, Perigolo N, Tripathi AK, Sprent JI, Young JP, James EK, et al. 2013. An invasive *Mimosa* in India does not adopt the symbionts of its native relatives. *Annals of Botany* **112**:179–196. DOI: <https://doi.org/10.1093/aob/mct112>, PMID: 23712450
- Gourion B**, Berrabah F, Ratet P, Stacey G. 2015. Rhizobium-legume symbioses: the crucial role of plant immunity. *Trends in Plant Science* **20**:186–194. DOI: <https://doi.org/10.1016/j.tplants.2014.11.008>, PMID: 25543258
- Hahn M**, Studer D. 1986. Competitiveness of a *nif*–bradyrhizobium japonicum mutant against the wild-type strain. *FEMS Microbiology Letters* **33**:143–148. DOI: <https://doi.org/10.1111/j.1574-6968.1986.tb01228.x>
- Hammerschmidt K**, Rose CJ, Kerr B, Rainey PB. 2014. Life cycles, fitness decoupling and the evolution of multicellularity. *Nature* **515**:75–79. DOI: <https://doi.org/10.1038/nature13884>, PMID: 25373677
- Hardin G**. 1968. The tragedy of the commons. *Science* **162**:1243–1248. DOI: <https://doi.org/10.1126/science.162.3859.1243>
- Heath KD**, Stinchcombe JR. 2014. Explaining mutualism variation: a new evolutionary paradox? *Evolution* **68**:309–317. DOI: <https://doi.org/10.1111/evo.12292>, PMID: 24303853
- Heath KD**, Tiffin P. 2009. Stabilizing mechanisms in a legume-rhizobium mutualism. *Evolution* **63**:652–662. DOI: <https://doi.org/10.1111/j.1558-5646.2008.00582.x>, PMID: 19087187
- Herre EA**, Knowlton N, Mueller UG, Rehner SA. 1999. The evolution of mutualisms: exploring the paths between conflict and cooperation. *Trends in Ecology & Evolution* **14**:49–53. DOI: [https://doi.org/10.1016/S0169-5347\(98\)01529-8](https://doi.org/10.1016/S0169-5347(98)01529-8), PMID: 10234251
- Hirsch AM**, Smith CA. 1987. Effects of *Rhizobium meliloti* *nif* and *fix* mutants on alfalfa root nodule development. *Journal of Bacteriology* **169**:1137–1146. DOI: <https://doi.org/10.1128/jb.169.3.1137-1146.1987>, PMID: 3818542
- Hoek TA**, Axelrod K, Biancalani T, Yurtsev EA, Liu J, Gore J. 2016. Resource availability modulates the cooperative and competitive nature of a microbial cross-feeding mutualism. *PLoS Biology* **14**:e1002540. DOI: <https://doi.org/10.1371/journal.pbio.1002540>, PMID: 27557335
- Johns NI**, Blazejewski T, Gomes AL, Wang HH. 2016. Principles for designing synthetic microbial communities. *Current Opinion in Microbiology* **31**:146–153. DOI: <https://doi.org/10.1016/j.mib.2016.03.010>, PMID: 27084981
- Jones EI**, Afkhami ME, Akçay E, Bronstein JL, Bshary R, Frederickson ME, Heath KD, Hoeksema JD, Ness JH, Pankey MS, Porter SS, Sachs JL, Scharnagl K, Friesen ML. 2015. Cheaters must prosper: reconciling theoretical and empirical perspectives on cheating in mutualism. *Ecology Letters* **18**:1270–1284. DOI: <https://doi.org/10.1111/ele.12507>
- Kawaharada Y**, Kelly S, Nielsen MW, Hjuler CT, Gysel K, Muszyński A, Carlson RW, Thygesen MB, Sandal N, Asmussen MH, Vinther M, Andersen SU, Krusell L, Thirup S, Jensen KJ, Ronson CW, Blaise M, Radutoiu S, Stougaard J. 2015. Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature* **523**:308–312. DOI: <https://doi.org/10.1038/nature14611>, PMID: 26153863
- Kiers ET**, Denison RF. 2008. Sanctions, cooperation, and the stability of plant-rhizosphere mutualisms. *Annual Review of Ecology, Evolution, and Systematics* **39**:215–236. DOI: <https://doi.org/10.1146/annurev.ecolsys.39.110707.173423>
- Kiers ET**, Rousseau RA, West SA, Denison RF. 2003. Host sanctions and the legume-rhizobium mutualism. *Nature* **425**:78–81. DOI: <https://doi.org/10.1038/nature01931>, PMID: 12955144
- Koch EJ**, Miyashiro T, McFall-Ngai MJ, Ruby EG. 2014. Features governing symbiont persistence in the squid-vibrio association. *Molecular Ecology* **23**:1624–1634. DOI: <https://doi.org/10.1111/mec.12474>, PMID: 24118200
- Ling J**, Zheng H, Katzianer DS, Wang H, Zhong Z, Zhu J. 2013. Applying reversible mutations of nodulation and nitrogen-fixation genes to study social cheating in *Rhizobium etli*-legume interaction. *PLoS One* **8**:e70138. DOI: <https://doi.org/10.1371/journal.pone.0070138>, PMID: 23922937
- Malik NS**, Calvert HE, Bauer WD. 1987. Nitrate induced regulation of nodule formation in soybean. *Plant Physiology* **84**:266–271. DOI: <https://doi.org/10.1104/pp.84.2.266>, PMID: 16665428
- Marchetti M**, Capela D, Glew M, Cruveiller S, Chane-Woon-Ming B, Gris C, Timmers T, Poinso V, Gilbert LB, Heeb P, Médigue C, Batut J, Masson-Boivin C. 2010. Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biology*, e1000280. DOI: <https://doi.org/10.1371/journal.pbio.1000280>, PMID: 20084095
- Marchetti M**, Catrice O, Batut J, Masson-Boivin C. 2011. *Cupriavidus taiwanensis* bacteroids in *Mimosa pudica* indeterminate nodules are not terminally differentiated. *Applied and Environmental Microbiology* **77**:2161–2164. DOI: <https://doi.org/10.1128/AEM.02358-10>, PMID: 21257807
- Marchetti M**, Clerissi C, Yousfi Y, Gris C, Bouchez O, Rocha E, Cruveiller S, Jauneau A, Capela D, Masson-Boivin C. 2017. Experimental evolution of rhizobia may lead to either extra- or intracellular symbiotic adaptation depending on the selection regime. *Molecular Ecology* **26**:1818–1831. DOI: <https://doi.org/10.1111/mec.13895>, PMID: 27770459

- Marchetti M**, Jauneau A, Capela D, Remigi P, Gris C, Batut J, Masson-Boivin C. 2014. Shaping bacterial symbiosis with legumes by experimental evolution. *Molecular Plant-Microbe Interactions* **27**:956–964. DOI: <https://doi.org/10.1094/MPMI-03-14-0083-R>, PMID: 25105803
- Marco DE**, Pérez-Arnedo R, Hidalgo-Perea Ángeles, Olivares J, Ruiz-Sainz JE, Sanjuán J. 2009. A mechanistic molecular test of the plant-sanction hypothesis in legume–rhizobia mutualism. *Acta Oecologica* **35**:664–667. DOI: <https://doi.org/10.1016/j.actao.2009.06.005>
- Masson-Boivin C**, Giraud E, Perret X, Batut J. 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends in Microbiology* **17**:458–466. DOI: <https://doi.org/10.1016/j.tim.2009.07.004>, PMID: 19766492
- Mueller UG**, Sachs JL. 2015. Engineering microbiomes to improve plant and animal health. *Trends in Microbiology* **23**:606–617. DOI: <https://doi.org/10.1016/j.tim.2015.07.009>, PMID: 26422463
- Nandasena KG**, O'hara GW, Tiwari RP, Howieson JG. 2006. Rapid in situ evolution of nodulating strains for *Biserrula pelecinus* L. through lateral transfer of a symbiosis island from the original mesorhizobial inoculant. *Applied and Environmental Microbiology* **72**:7365–7367. DOI: <https://doi.org/10.1128/AEM.00889-06>, PMID: 16936054
- Nandasena KG**, O'Hara GW, Tiwari RP, Sezmiş E, Howieson JG. 2007. In situ lateral transfer of symbiosis islands results in rapid evolution of diverse competitive strains of mesorhizobia suboptimal in symbiotic nitrogen fixation on the pasture legume *Biserrula pelecinus* L. *Environmental Microbiology* **9**:2496–2511. DOI: <https://doi.org/10.1111/j.1462-2920.2007.01368.x>, PMID: 17803775
- Oono R**, Anderson CG, Denison RF. 2011. Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. *Proceedings of the Royal Society B: Biological Sciences* **278**:2698–2703. DOI: <https://doi.org/10.1098/rspb.2010.2193>
- Oono R**, Denison RF, Kiers ET. 2009. Controlling the reproductive fate of rhizobia: how universal are legume sanctions? *New Phytologist* **183**:967–979. DOI: <https://doi.org/10.1111/j.1469-8137.2009.02941.x>, PMID: 19594691
- Papkou A**, Gokhale CS, Traulsen A, Schulenburg H. 2016. Host-parasite coevolution: why changing population size matters. *Zoology* **119**:330–338. DOI: <https://doi.org/10.1016/j.zool.2016.02.001>, PMID: 27161157
- Perret X**, Staehelin C, Broughton WJ. 2000. Molecular basis of symbiotic promiscuity. *Microbiology and Molecular Biology Reviews* **64**:180–201. DOI: <https://doi.org/10.1128/MMBR.64.1.180-201.2000>, PMID: 10704479
- Porter SS**, Simms EL. 2014. Selection for cheating across disparate environments in the legume-rhizobium mutualism. *Ecology Letters* **17**:1121–1129. DOI: <https://doi.org/10.1111/ele.12318>, PMID: 25039752
- R Core Team**. 2014. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Radutoiu S**, Madsen LH, Madsen EB, Felle HH, Umehara Y, Grønlund M, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J. 2003. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425**:585–592. DOI: <https://doi.org/10.1038/nature02039>, PMID: 14534578
- Regus JU**, Quides KW, O'Neill MR, Suzuki R, Savory EA, Chang JH, Sachs JL. 2017. Cell autonomous sanctions in legumes target ineffective rhizobia in nodules with mixed infections. *American Journal of Botany* **104**:1299–1312. DOI: <https://doi.org/10.3732/ajb.1700165>
- Remigi P**, Zhu J, Young JPW, Masson-Boivin C. 2016. Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. *Trends in Microbiology* **24**:63–75. DOI: <https://doi.org/10.1016/j.tim.2015.10.007>, PMID: 26612499
- Sachs JL**, Ehinger MO, Simms EL. 2010. Origins of cheating and loss of symbiosis in wild *Bradyrhizobium*. *Journal of Evolutionary Biology* **23**:1075–1089. DOI: <https://doi.org/10.1111/j.1420-9101.2010.01980.x>, PMID: 20345811
- Sachs JL**, Mueller UG, Wilcox TP, Bull JJ. 2004. The evolution of cooperation. *The Quarterly Review of Biology* **79**:135–160. DOI: <https://doi.org/10.1086/383541>, PMID: 15232949
- Shou W**. 2015. Acknowledging selection at sub-organismal levels resolves controversy on pro-cooperation mechanisms. *eLife* **4**:e10106. DOI: <https://doi.org/10.7554/eLife.10106>, PMID: 26714105
- Tarnita CE**. 2017. The ecology and evolution of social behavior in microbes. *The Journal of Experimental Biology* **220**:18–24. DOI: <https://doi.org/10.1242/jeb.145631>, PMID: 28057824
- Thies JE**, Woomer PL, Singleton PW. 1995. Enrichment of bradyrhizobium spp populations in soil due to cropping of the homologous host legume. *Soil Biology and Biochemistry* **27**:633–636. DOI: [https://doi.org/10.1016/0038-0717\(95\)98643-3](https://doi.org/10.1016/0038-0717(95)98643-3)
- van Noorden GE**, Verbeek R, Dinh QD, Jin J, Green A, Ng JL, Mathesius U. 2016. Molecular signals controlling the inhibition of nodulation by nitrate in medicago truncatula. *International Journal of Molecular Sciences* **17**:E1060. DOI: <https://doi.org/10.3390/ijms17071060>, PMID: 27384556
- Vigneron A**, Masson F, Vallier A, Balmand S, Rey M, Vincent-Monégat C, Aksoy E, Aubailly-Giraud E, Zaidman-Rémy A, Heddi A. 2014. Insects recycle endosymbionts when the benefit is over. *Current Biology* **24**:2267–2273. DOI: <https://doi.org/10.1016/j.cub.2014.07.065>, PMID: 25242028
- Virta M**, Lineri S, Kankaanpää P, Karp M, Peltonen K, Nuutila J, Lilius EM. 1998. Determination of complement-mediated killing of bacteria by viability staining and bioluminescence. *Applied and Environmental Microbiology* **64**:515–519. PMID: 9464386
- Visick KL**, McFall-Ngai MJ. 2000. An exclusive contract: specificity in the *Vibrio fischeri*-*Euprymna scolopes* partnership. *Journal of Bacteriology* **182**:1779–1787. DOI: <https://doi.org/10.1128/JB.182.7.1779-1787.2000>, PMID: 10714980

- Werner GDA**, Cornwell WK, Sprent JI, Kattge J, Kiers ET. 2014. A single evolutionary innovation drives the deep evolution of symbiotic N₂-fixation in angiosperms. *Nature Communications* **5**:5087. DOI: <https://doi.org/10.1038/ncomms5087>
- Westhoek A**, Field E, Rehling F, Mulley G, Webb I, Poole PS, Turnbull LA. 2017. Policing the legume-Rhizobium symbiosis: a critical test of partner choice. *Scientific Reports* **7**:1419. DOI: <https://doi.org/10.1038/s41598-017-01634-2>, PMID: 28469244

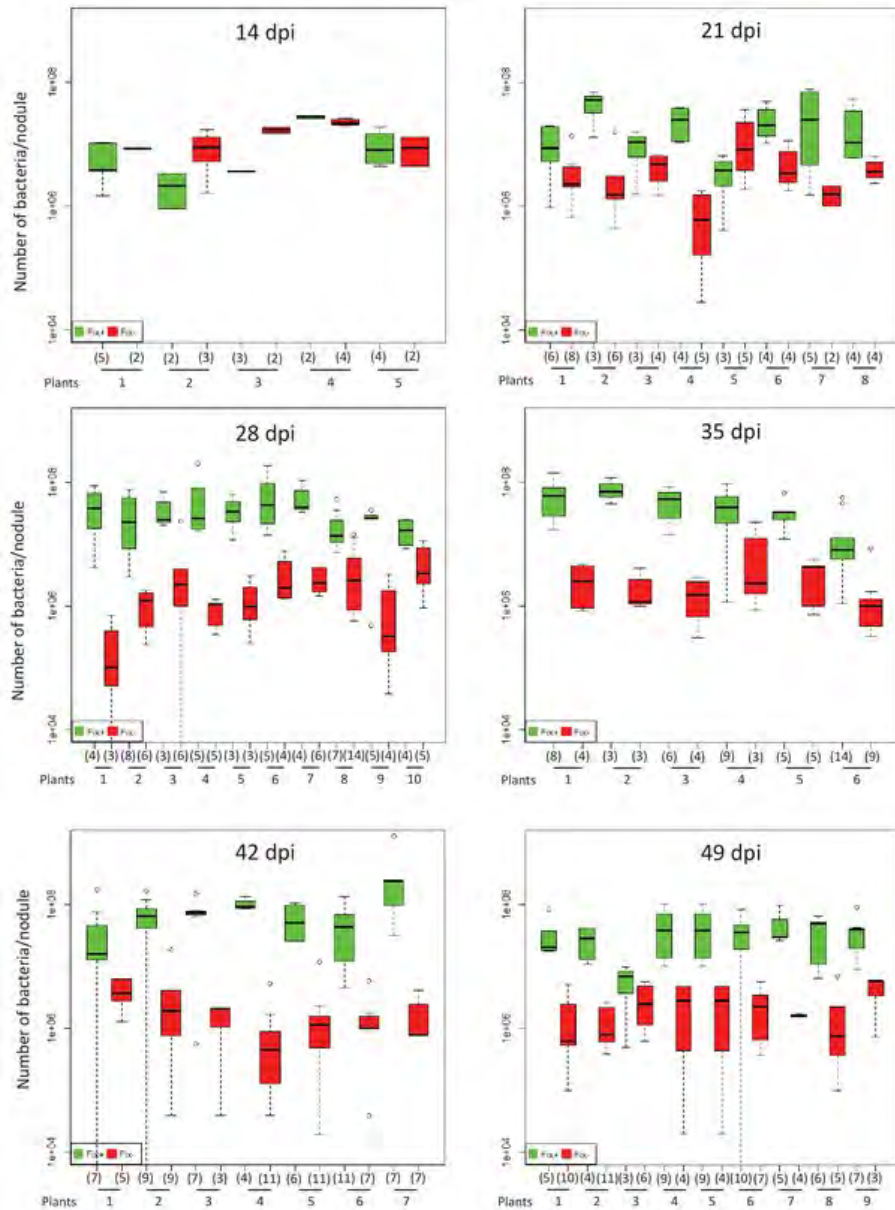
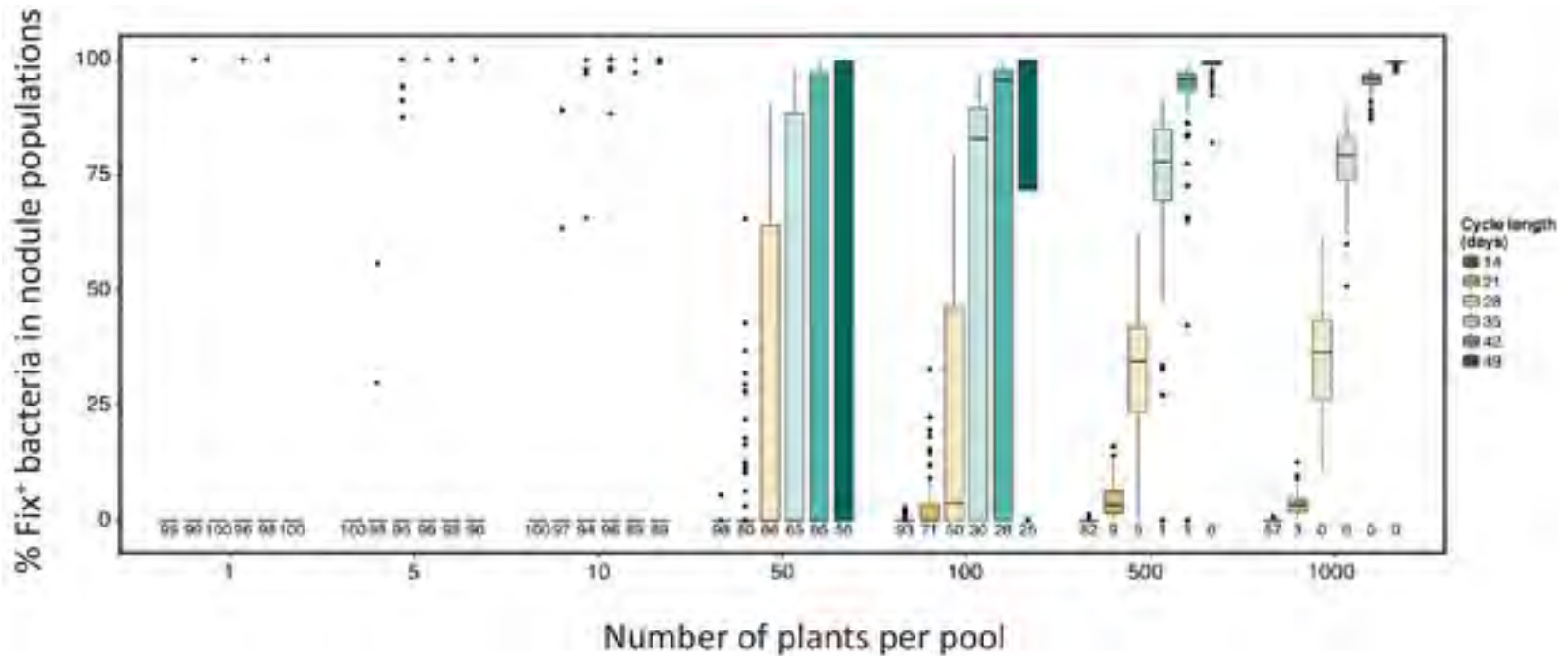


Figure 1 – figure supplement 1. Kinetics of reproductive fitness of Fix+ and Fix- bacteria in nodules following co-inoculation of *M. pudica*. *M. pudica* plants were co-inoculated with a mixture of Fix+ and Fix- strains at a 1/1 ratio. Nodules were individually analyzed by plating their bacterial population. Box plots represent the number of bacteria per nodule for each individual plant at each time points. Central rectangles span the first quartile to the third quartile (that is, the interquartile range or IQR), bold segments inside rectangles show the median, unfilled circles indicate suspected outliers, whiskers above and below the box show either the locations of the minimum and maximum in the absence of suspected outlying data or 1.5 X IQR if an outlier is present.

Figure 9 – figure supplement 1. Effect of cycle length and plant numbers on the predicted distributions of Fix+ population sizes. Model simulations were performed with an initial proportion of 1% Fix+ in the bacterial population inoculated to a pool of plants. The length of each cycle and the number of plants per pool varied as indicated in the legend. Boxplots represent the distribution of the final proportion of Fix+ clones from 100 simulations. The length of inoculation cycles ranged from 14 to 49 days and the number of plants per pool from 1 to 1000. Numbers underneath each boxplot indicate the number of replicate simulations where Fix+ sub-populations became extinct after four cycles.



Supplementary File 1. Primers used in this study.

Primer name	Primer sequence 5'-3'	Use
oCBM1821	CGGGATCCGATCACACCCAACGATGAGA	Amplification of the <i>nifH</i> upstream region
oCBM2362	GCTCTAGAGTTGCCAAGCGACGTATTT	Amplification of the <i>nifH</i> upstream region
oCBM1822	CGGGATCCGTGCAGATCGACGAGTCAA	Amplification of the <i>nifH</i> downstream region
oCBM2363	CCGGAATTCGAGTACTGGCCACAGCCAAC	Amplification of the <i>nifH</i> downstream region
oCBM1824	ACGAAATTGTCCGATGGCTA	Verification of the <i>nifH</i> deletion
oCBM1825	TAGTAATTGCGACGCGATCC	Verification of the <i>nifH</i> deletion
oCBM2619	TCTAGAGTGGTGGTGTATCTCGCAATC	Amplification of the <i>glnS</i> -Ralta_A206 upstream region
oCBM2620	GGTACCCTAGGGTGAGGCAAGGCGTTCAG	Amplification of the <i>glnS</i> -Ralta_A0206 upstream region
oCBM2621	GGTACCACTAGTCTGAACGCCTTGCTCAC	Amplification of the <i>glnS</i> -Ralta_A0206 downstream region
oCBM2622	GAATTCAGCTGCATCACTTCGTCAC	Amplification of the <i>glnS</i> -Ralta_A0206 downstream region

Chapitre III: Identification d'un nouveau gène *nod* impliqué dans une structure atypique des facteurs Nod

III.1. Introduction

A l'issue des 16 cycles d'évolution, les 3 lignées longues dérivant de l'ancêtre nodulant et extracellulaire CBM356 ne semblaient pas avoir acquis de capacité à l'intracellularité (Marchetti et al. 2016). En effet, aucun des 3 clones isolés des populations finales des lignées P, R et T n'était capable d'infecter les cellules des nodules. Cela indiquait que, soit aucune mutation adaptative pour l'infection intracellulaire n'était apparue dans les clones des lignées P, R et T, soit que l'effet de mutations adaptatives pour l'infection était annulé par une autre ou plusieurs autres mutations.

Les travaux de l'équipe préalablement à mon arrivée avaient montré que dans le cas de la lignée R, il s'agissait du deuxième cas de figure, et avaient permis d'identifier une mutation antagoniste candidate. En effet, le séquençage de R16 avait révélé la présence d'une mutation non synonyme affectant le régulateur HrpG (A179V), dont l'inactivation par une mutation stop avait été montrée adaptative pour l'infection (Marchetti et al. 2010). La reconstruction de cette mutation dans l'ancêtre CBM356 conférait à cette souche la capacité à infecter –de façon assez rudimentaire- les cellules des nodules formés, confirmant son caractère adaptatif (**Figure III.1.1**) (Marchetti et al. 2016). L'incapacité de R16 à infecter les cellules des nodules indiquait donc l'existence d'une ou plusieurs autres modifications du génome de R16 qui contrebalançaient les effets de la mutation *hrpGA179V*. Le remplacement du pRalta évolué de R16 par le pRalta original avait confirmé cette hypothèse. En effet, ce remplacement restaurait le phénotype infectieux obtenu avec la souche CBM356*hrpGA179V* (**Figure III.1.1**) (Marchetti et al. 2016). L'analyse du pRalta de R16 avait montré que le plasmide de ce clone évolué portait une large délétion de 50194 bp (**Figure III.1.2**). Cette délétion supprimait 62 gènes encodant majoritairement des transposases, mais aussi des protéines de fonction inconnue ou des protéines putatives, dont une stérol désaturase, produit du gène *pRalta_0471*. La complémentation de *pRalta_0471* dans le clone évolué R16 avait restauré le phénotype infectieux de R16 pRalta, suggérant que ce gène était requis pour l'infection intracellulaire (**Figure III.1.3**).

Pour cette seconde partie de ma thèse, j'ai donc analysé le rôle de ce gène, que nous avons renommé *noeM*, dans le système de référence *C. taiwanensis*/*M. pudica*. La position de ce gène en aval d'une *nod*-box laissait suspecter un rôle dans la structure des Facteurs Nod, ce que nous avons montré. Par contre nous n'avons pas pu mettre en évidence de rôle de ce gène dans l'infection intracellulaire chez *C. taiwanensis*. Les résultats de ces travaux sont décrits dans l'article Daubech et al., 2019, *noeM*, a new nodulation gene involved in the biosynthesis of Nod factors with an open-chain oxidized terminal residue and in the symbiosis with *Mimosa pudica*, publié dans la revue *Molecular Plant-Microbe Interactions* (MPMI).

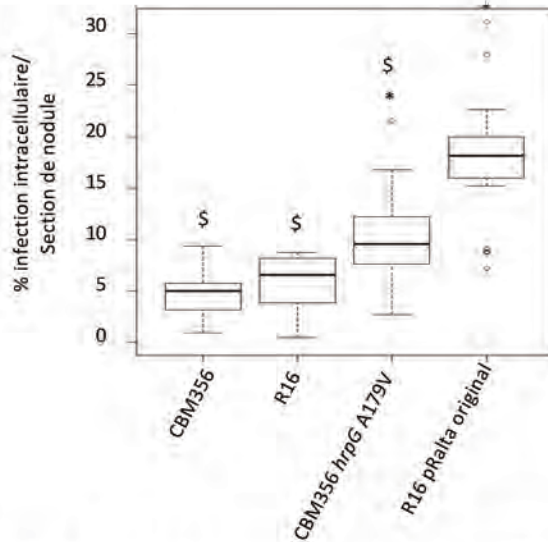


Figure III.1.1: Acquisition de l'infection intracellulaire dans la lignée R. La quantification de l'infection intracellulaire a été obtenue à partir de 12 à 23 nodules différents formés par CBM356, R16, CBM356hrpGA179V et R16 complété avec le pRalta original. Les rectangles centraux s'étendent du premier quartile au troisième quartile (c'est-à-dire l'intervalle interquartile ou II), les segments en gras à l'intérieur des rectangles indiquent la médiane, les cercles non remplis indiquent les valeurs aberrantes présumées, les moustaches au-dessus et en dessous de la boîte indiquent les emplacements du minimum et du maximum en l'absence de données aberrantes suspectées ou 1,5X II si une valeur aberrante est présente. * et § indiquent les différences significatives avec l'ancêtre CBM356 et la souche R16, respectivement ($P < 0.05$; Kruskal-Wallis test) (adapté de Marchetti *et al*, 2016).



Figure III.1.2: Position de la délétion du pRalta incluant *noeM* chez le clone R16. Le séquençage du clone R16 a permis de mettre une large délétion de 50194 bp sur le pRalta en position 357254-407448. Le gène *pRalta_0471/noeM* est encerclé en vert.

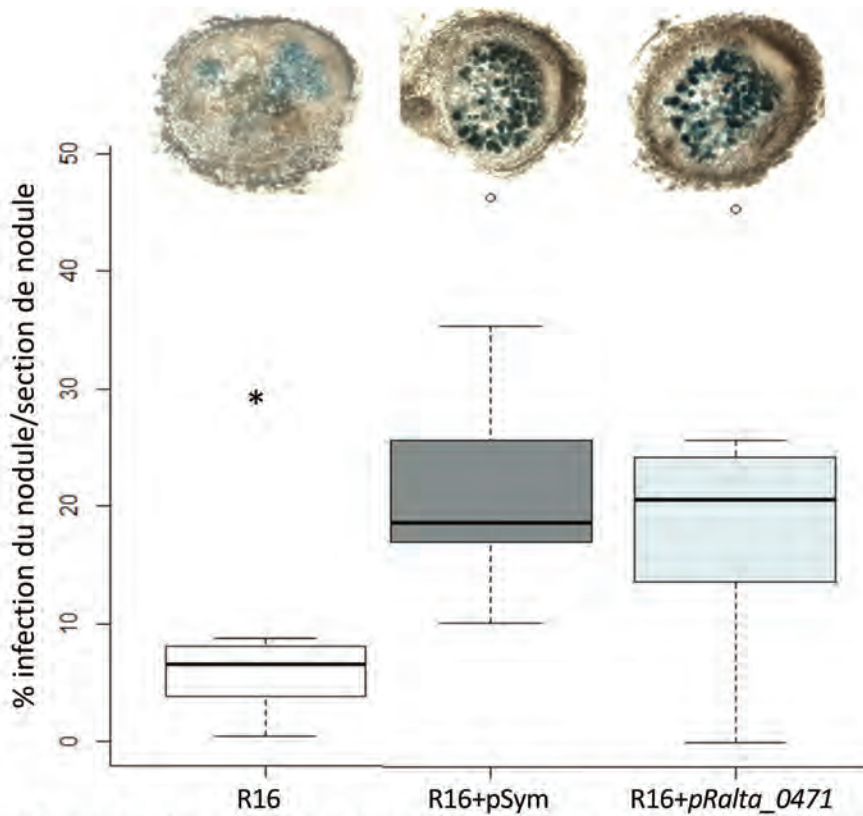


Figure III.1.3: Capacité d'infection du clone R16 et de ses complémentations avec le pRalta d'origine et le gène *pRalta_0471*. La quantification de l'infection intracellulaire a été obtenue à partir de 12 à 22 nodules différents. Les rectangles centraux s'étendent du premier quartile au troisième quartile (c'est-à-dire l'intervalle interquartile ou I), les segments en gras à l'intérieur des rectangles indiquent la médiane, les moustaches au-dessus et en dessous de la boîte indiquent les emplacements du minimum et du maximum. * indiquent les différences significatives avec la souche R16 ($P < 0.05$; Kruskal-Wallis test).

III.2. Article

Ma contribution dans ce travail a été de construire toutes les souches de *C. taiwanensis* utilisées. J'ai également réalisé les tests de compétition pour la nodulation et fitness *in planta* avec *C. taiwanensis*, les mesures de cinétique de nodulation et d'expression des gènes ainsi que les mesures d'expression des gènes par qRT-PCR. J'ai aussi réalisé l'analyse phylogénétique de NoeM/*noeM* et préparé les cultures bactériennes utilisées pour les analyses de la structure des facteurs Nod. Enfin, j'ai contribué à l'écriture et la relecture de la publication ainsi qu'à la production des figures.

noeM, a New Nodulation Gene Involved in the Biosynthesis of Nod Factors with an Open-Chain Oxidized Terminal Residue and in the Symbiosis with *Mimosa pudica*

Benoit Daubech,¹ Verena Poinsoot,² Agnieszka Klonowska,³ Delphine Capela,¹ Clémence Chaintreuil,⁴ Lionel Moulin,³ Marta Marchetti,¹ and Catherine Masson-Boivin^{1,†}

¹ LIPM, Université de Toulouse, INRA, CNRS, Castanet-Tolosan, France

² Université de Toulouse 3, UPS CNRS 5623, UMR, Lab IMRCP, F-31062 Toulouse, France

³ IRD, CIRAD, Université Montpellier, IPME, Montpellier, France

⁴ Université Montpellier, CIRAD, INRA, IRD, Montpellier SupAgro, LSTM, Montpellier, France

Accepted 28 July 2019.

The β -rhizobium *Cupriavidus taiwanensis* is a nitrogen-fixing symbiont of *Mimosa pudica*. Nod factors produced by this species were previously found to be pentameric chitin-oligomers carrying common C18:1 or C16:0 fatty acyl chains, N-methylated and C-6 carbamoylated on the nonreducing terminal N-acetylglucosamine and sulfated on the reducing terminal residue. Here, we report that, in addition, *C. taiwanensis* LMG19424 produces molecules where the reducing sugar is open and oxidized. We identified a novel nodulation gene located on the symbiotic plasmid pRalta, called *noeM*, which is involved in this atypical Nod factor structure. *noeM* encodes a transmembrane protein bearing a fatty acid hydroxylase domain. This gene is expressed during symbiosis with *M. pudica* and requires NodD and luteolin for optimal expression. The closest *noeM* homologs formed a separate phylogenetic clade containing rhizobial genes only, which are located on symbiosis plasmids downstream from a *nod box*. Corresponding proteins, referred to as NoeM, may have specialized in symbiosis via the connection to the nodulation pathway and the spread in rhizobia. *noeM* was mostly found in isolates of the Mimoseae tribe, and specifically detected in all tested strains able to nodulate *M. pudica*. A *noeM* deletion mutant of *C. taiwanensis* was affected for the nodulation of *M. pudica*, confirming the role of *noeM* in the symbiosis with this legume.

Keywords: *Cupriavidus taiwanensis*, host specificity, *Mimosa pudica*, Nod factors, nodulation, rhizobium-legume symbiosis

[†]Corresponding author: C. Masson-Boivin; catherine.masson@inra.fr

Funding: This work was supported by funds from the French National Research Agency (ANR-16-CE20-0011-01) and the French Laboratory of Excellence project “TULIP” (ANR-10-LABX-41; ANR-11-IDEX-0002-02). B. Daubech was supported by an INRA-Région Occitanie fellowship.

*The e-Xtra logo stands for “electronic extra” and indicates that six supplementary figures and three supplementary tables are published online.

The author(s) declare no conflict of interest.

© 2019 The American Phytopathological Society

Soil bacteria collectively referred to as rhizobia are able to form nitrogen-fixing symbioses with leguminous plants. These symbioses are manifested by the formation of novel root (or occasionally stem) organs called nodules, within which bacteria are internalized and fix nitrogen to the benefit of the plant. Most of the rhizobia achieve nodulation by using the so-called Nod strategy, although two alternative strategies—the type three secretion system (T3SS) and the genetically uncharacterized non Nod/non-T3SS strategies—have been identified (Masson-Boivin and Sachs 2018). In the Nod strategy, host plants excrete inducing compounds—mainly flavonoids—that activate LysR-type transcriptional regulator NodD proteins of compatible rhizobia (Mulligan and Long 1985; Nelson and Sadowsky 2015). NodD proteins bind to conserved regulatory DNA sequences, termed *nod* boxes, which are located in the promoters of nodulation (*nod*) genes (Rostas et al. 1986), resulting in their transcriptional activation. Flavonoids binding to NodD enhance *nod* gene expression (Chen et al. 2005; Peck et al. 2006). The *nod* genes encode enzymes that synthesize signaling lipo-chitooligosaccharides (LCOs) called Nod factors (NFs) (Downie 1998) that are perceived by plant receptors (Broghammer et al. 2012; Madsen et al. 2003) and activate a host genetic program governing the nodulation and infection processes (Gough and Cullimore 2011; Oldroyd et al. 2011). NFs are chitin oligomers of β -(1,4)-linked N-acetyl-D-glucosaminyl (GlcNAc) residues (two to six units) that are N-acetylated on their nonreducing end (D’Haeze and Holsters 2002). The NF structure differs in the number of GlcNAc residues of the chitin oligosaccharide backbone—generally three, four, or five—in the nature of the fatty acyl group and in the substituents at the—generally terminal—residues. A given strain often produces a variety of NF compounds. These structural variations, which are strain or species specific, are involved in host specificity (Lerouge et al. 1990; Schultze et al. 1992) together with the range of flavonoids excreted by the plant (Liu and Murray 2016). The core lipo-chitooligosaccharidic NF structure is encoded by the common *nodABC* genes, while other *nod*, *nol*, or *noe* genes encode modifications of this structure (Perret et al. 2000). The essential *nod* genes are clustered together with nitrogen-fixation genes in mobile genetic elements (MGE) such as symbiotic plasmids or genomic islands. During evolution, these symbiotic MGE have spread to different species, converting soil bacteria into mutualistic symbionts of legumes (Nandasena et al. 2006;

Sullivan et al. 1995), a process that can be partly or fully reproduced in the lab (Marchetti et al. 2010, 2014; Sullivan and Ronson 1998). Consequently, extant rhizobia are phylogenetically diverse bacteria that belong to hundreds of species distributed in at least 14 genera of α - and β -proteobacteria, referred to as α - and β -rhizobia (Remigi et al. 2016).

The β -rhizobium *Cupriavidus taiwanensis* LMG19424 is a specific symbiont of *Mimosa* spp., including *Mimosa pudica* (Chen et al. 2003b). *C. taiwanensis* enters *M. pudica* roots via infection threads that ultimately release bacteria into nodule cells, where bacteroids are not terminally differentiated (Marchetti et al. 2011). Genome sequencing revealed that this strain possesses a large symbiotic plasmid, pRalta, harboring one *nod* locus (Amadou et al. 2008). This locus contains a *nodBCIJHASUQ* operon preceded by a *nod* box (*nod* box1) as well as one *nodD* gene divergently transcribed. NFs produced by *C. taiwanensis* were found to be pentameric chitin-oligomers and N-acylated by vaccenic acid (C18:1) or palmitic acid (C16:0) on their nonreducing end (Amadou et al. 2008). Consistent with the presence of *nodU*, *nodS*, and *nodH* that are responsible for the addition of O-carbamoyl, N-methyl, and sulfate (S) groups, respectively, on the core NF, most of the molecules are sulfated at the reducing end and substituted by an N-methyl and a carbamoyl group at the terminal nonreducing sugar. A second *nod* box (*nod* box2) has been predicted 92 kb from the *nod* locus (Amadou et al. 2008). This *nod* box is upstream from a single gene, *pRalta_0471*, that has been annotated as a fatty acid hydrolase. This gene was found in all sequenced *C. taiwanensis* strains (Clerissi et al. 2018) and to have orthologs in other recently sequenced *M. pudica* symbionts, *Paraburkholderia phymatum* STM815 (BURPHP2_0638), and *Rhizobium mesoamericanum* STM3625 (BN77v2_p2140043) (Moulin et al. 2013, 2014). It has been shown to be induced by *M. pudica* exudates in all three species (Klonowska et al. 2018).

In this study, we characterize *pRalta_0471*, which we renamed *noeM*. We show that *noeM* is involved in the synthesis of atypical NF molecules and plays a role in the nodulation of *M. pudica*. The phylogenetic distribution of *noeM* supports its role in the symbiosis with *Mimosa* spp.

RESULTS

noeM is expressed during symbiosis and requires NodD and luteolin for optimal expression.

To analyze *noeM* regulation, we monitored the expression of *C. taiwanensis noeM* (*pRalta_0471*) in both free-living and symbiotic conditions. For that, we constructed a chromosomal *noeM*-mCherry fusion in *C. taiwanensis* CBM832, a streptomycin-resistant derivative of the wild-type LMG19424 strain. This was done by cloning a 404-bp fragment containing *nod* box2 upstream *noeM* fused to the mCherry gene at the *glmS* intergenic region of CBM832. Chromosomal mCherry fusions with the promoters of the nitrogen-fixation *nifH* and the constitutively expressed chloroplastic *psbA* (*Pps*) genes were used as controls.

To check whether *noeM* is expressed during the symbiotic process, seedlings of *M. pudica* were single inoculated with CBM832 bearing either the *noeM*, *psbA*, or *nifH* promoter fusion. Plants were harvested at different time points after inoculation and examined using confocal laser microscopy and light microscopy. Unlike *nifH*, which is only expressed in nodule cells (Marchetti et al. 2011), *noeM* was expressed at a level similar to that of the constitutively expressed *psbA* gene at all stages of symbiosis (i.e., in the rhizosphere, in the infection sites, in the infection threads, and within bacteroids) (Fig. 1).

The *nod* genes of *C. taiwanensis* were shown to be induced by the flavonoid luteolin (Amadou et al. 2008), likely through

the activation of the single *nodD* regulatory gene identified in the LMG19424 genome. To analyze the effect of luteolin and *nodD* on *noeM* expression in vitro, we constructed a *nodD* deletion mutant of the CBM832 strain carrying the *noeM*-mCherry fusion. The expression of *noeM* and *nodB* in CBM832 and CBM832 Δ *nodD* was first measured by quantitative reverse-transcription (RT)-PCR in cultures grown in minimal medium with or without luteolin (10 μ M) for 5 h 30 min (Fig. 2). The expression of *nodB* and *noeM* increased by a factor of

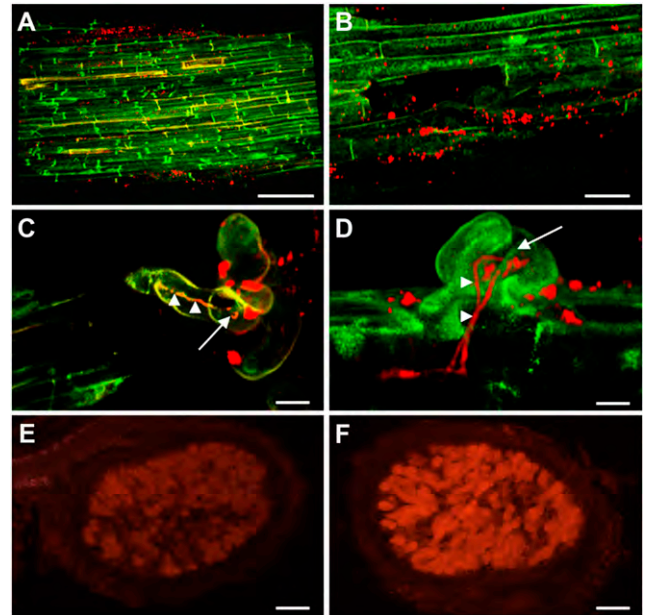


Fig. 1. *noeM* expression during symbiosis. Imaging of *Mimosa pudica* plants inoculated with *Cupriavidus taiwanensis* bearing a chromosomal **A**, **C**, and **E**, *PpsbA*-mCherry or **B**, **D**, and **F**, *pnoeM*-mCherry fusion at **A** and **B**, 24; **C** and **D**, 7; or **E** and **F**, 21 days postinoculation. Red indicates mCherry, green indicates autofluorescence of plant cells, arrow indicates infection site, and arrowhead indicates infection thread. Scale bars: 100 μ m (**A**, **B**, **E**, and **F**) and 10 μ m (**C** and **D**).

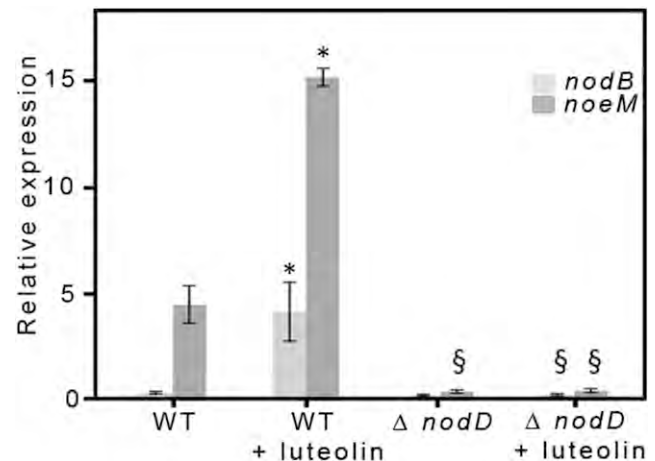


Fig. 2. Relative expression of the *nodB* and *noeM* genes in wild-type (WT) and mutant (Δ *nodD*) strains of *Cupriavidus taiwanensis*. Strains were cultivated in minimal medium with or without 10 μ M luteolin. The expression of *nodB* and *noeM* were quantified by quantitative PCR 5 h 30 min after luteolin induction. Raw expression levels were normalized by the *hisB* and *rpoA* gene expressions. Symbols: * indicates statistically significant difference between the treatment with or without luteolin for the same strain ($P < 0.05$, Student's *t* test with paired data) and § indicates statistically significant difference between the WT and mutant strain for the same treatment ($P < 0.05$, Student's *t* test with paired data).

approximately 14.5 and 3.4, respectively, in the presence of luteolin. No or very weak expression was measured in the CBM832 Δ *nodD* background. In addition, we monitored the expression of *noeM* in CBM832 and CBM832 Δ *nodD* grown in microplates in minimal medium with or without luteolin by measuring the optical density (OD) and mCherry fluorescence of the cultures over 24 h (Supplementary Fig. S1). In the absence of luteolin, the kinetics of *noeM* expression was similar to that of the constitutively expressed *psbA* fusion, although weaker. Its expression started increasing 5 h after medium inoculation and reached a plateau at approximately 17 h. *noeM* expression was significantly higher in the presence of luteolin from time = 17 h. By contrast, *noeM* was weakly expressed in a Δ *nodD* background, although higher than *nifH* expression, whatever the presence of luteolin.

Altogether, these data indicated that *noeM* was induced by luteolin, the inducer of *nod* genes in *C. taiwanensis*, and requires NodD for basal and luteolin-induced expression.

noeM is involved in NF biosynthesis.

Because *noeM* is predicted to belong to the fatty acid hydroxylase family of genes, we investigated the possibility that *noeM* is involved in the synthesis of an additional NF molecule—possibly by modifying the NF lipid moiety—that has been missed in the previous study of *C. taiwanensis* NFs (Amadou et al. 2008).

We first reanalyzed NFs produced by *C. taiwanensis* LMG19424 and its streptomycin-resistant derivative CBM832 in response to luteolin activation, taking advantage of technical advances in mass spectrometry (MS) technology that allow an increased sensitivity of the online liquid chromatography tandem mass spectrometry (MS/MS) coupling. Ultrahigh-performance liquid chromatography (UPLC)-QTrap-enhanced MS in the positive mode analyses showed that, in addition to the already identified compounds (Fig. 3A), the strains produce NF molecules that were not described before. Enhanced product ion scan MS/MS experiments in the positive mode were performed on the two major compounds, *m/z* 1,393.6 (formerly described) and *m/z* 1,423.7 (the undescribed one) (Fig. 4A). The MS/MS spectra recorded for these two compounds exhibited the same fragmentation pattern (the same B1, B2, B3, and B4) (Supplementary Fig. S2). This clearly indicated that the modification producing the mass shift of 30 atomic mass units (amu) (1,423 to 1,393) specifically affects the fifth GlcNAc (“reducing-end”), while it still carries an S group.

Only a few sugar modifications result in a mass difference of 30 amu: addition of a CH₂O group or addition of two oxygen with loss of two hydrogen (+2O -2H atoms). To determine which modification is involved, the NF fractions were submitted to Fourier-transform ion cyclotron resonance MS/MS analyses in the positive and in the negative electrospray ionization (ESI) modes (resolution of 195,000). The major member of the new family exhibited an exact mass of 1,421.5857. The mass defect (0.5857) and the isotopic profile (*M* + 1 = 67%) were compatible with the NF generic structure. The MS/MS fragmentation pattern obtained in the negative mode indicated clearly the loss of CO₂ (-43.99 amu) starting from the molecular ion 1,421.59. The high-resolution *m/z* values of the molecular ions of the classical (1,391.6115) and the novel (1,421.5857) structures were compared. The exact mass difference was 29.974 amu between the two NF series. The gain of a CH₂O would result in a mass difference of 30.011 amu. The gain of two O associated with the loss of two H would result in a difference of 29.988 amu. With a resolution of 195,000, the third decimal is significant at a mass of 30 amu. Therefore, the modification here observed on the classical NF structure is an oxidation of the terminal sugar unit. To explain this +2O -2H,

the only possibility is that the reducing sugar is open and oxidized into a diacid. This implies that the S group located on the O-6 for the common structure has to migrate to the neighbor O-5. The presence of the carboxylic groups explains the loss of CO₂ observed on the MS/MS spectrum.

To assess the presence of the carboxylic groups, we esterified the sample with diazomethane and reanalyzed them by ESI MS/MS. The new characterized NF-like molecules were demonstrated as esterified twice with diazomethane. This confirmed the presence of two free carboxylic acids (*m/z* 1,437.6 and *m/z* 1,451.7).

Yet, the arrangement of these chemical groups (i.e., two COOH, one N-acetyl, and one SO₃H) remained hypothetical. We were not able to write any alternative formula other than a lipochito-tetramer decorated with a sulfated, dehydroxylated, and N-acetylated adipic side chain (LCO-V open-chain oxidized) (Fig. 3B). The newly described NF (MW 1,422.6) is a member of a family that corresponds well to the classic family. The structural variations mostly affected the lipid chain and the presence of the carbamoyl group (LCO-V open-chain oxidized C18:1/C18:0/C16:0, NMe/H, Cb/H, and S) (Fig. 3B; Table 1; Supplementary Table S1). In both NF molecules, the classical and open-oxidized ones, the acyl chain was mainly a C18:1, the C18:0 and C16:0 being present in less than 10% of the molecules. The percentage of open-chain oxidized structures in total *C. taiwanensis* NF compounds varied from 38 to 77% in our culture conditions (Table 1).

We constructed a *noeM* deletion mutant of CBM832, and determined the structure of the NFs it produced. NFs secreted by CBM832 Δ *noeM* were found to bear the same acyl groups as

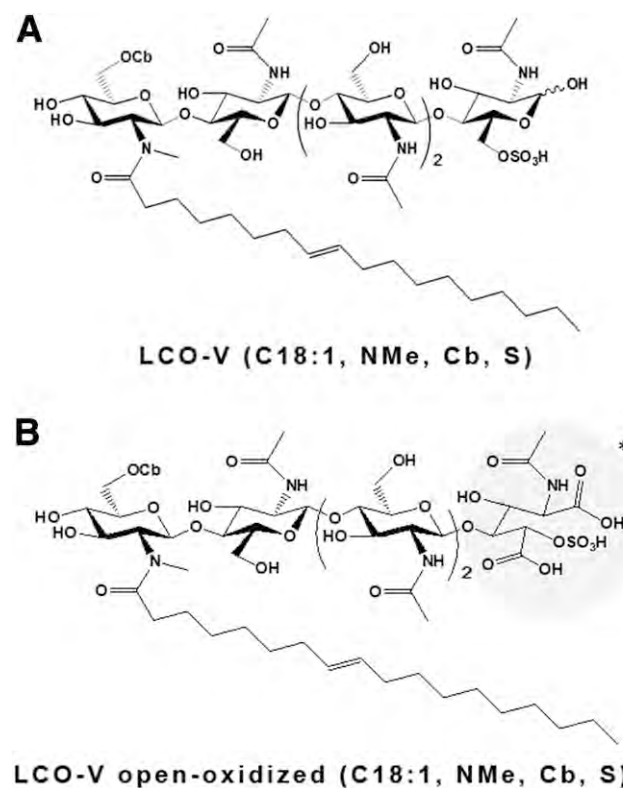


Fig. 3. Chemical structure of main Nod factors produced by *Cupriavidus taiwanensis*. **A**, Molecules previously identified and **B**, molecules newly identified in this study. LCO = lipo-chitoooligosaccharide. An asterisk (*) indicates the open-oxidized terminal residue, where the position of the two COOH, the N-acetyl, and the SO₃H groups is hypothetical. In 10% of the classical or open-oxidized molecules, the C18:1 acyl chain is replaced by a C18:0 or C16:0 group and the carbamoyl or N-Me can be absent.

the wild-type strain. However, CBM832 Δ *noeM* only produced the usual LCO-V (C18:1/C18:0/C16:0, NMe/H, Cb/H, and S) structure when the open oxidized family is completely absent (Fig. 4B; Table 1). The complementation of CBM832 Δ *noeM* with the original *noeM* gene restored the production of LCO-V open-chain oxidized (C18:1/C18:0/C16:0, NMe/H, Cb/H, and S) molecules (Table 1). To confirm the role of *noeM* using another genomic background, we introduced the wild-type or the *noeM*-deleted pRalta in *R. solanacearum* GMI1000 and analyzed the NFs of the two strains. GMI1000pRalta produced

both classical and open-chain oxidized LCO-V structures, while GMI1000pRalta Δ *noeM* produced only the classical structure (Table 1). This indicated that *noeM* is a *nod* gene involved in the biosynthesis of an atypical NF structure.

noeM has specialized in rhizobial symbiosis.

To analyze the phylogeny of *noeM*, we searched for homologs of pRalta_0471 in data banks. We retrieved 78 different NoeM homologs originating from 136 different strains belonging to α - and β -proteobacteria (Supplementary Table

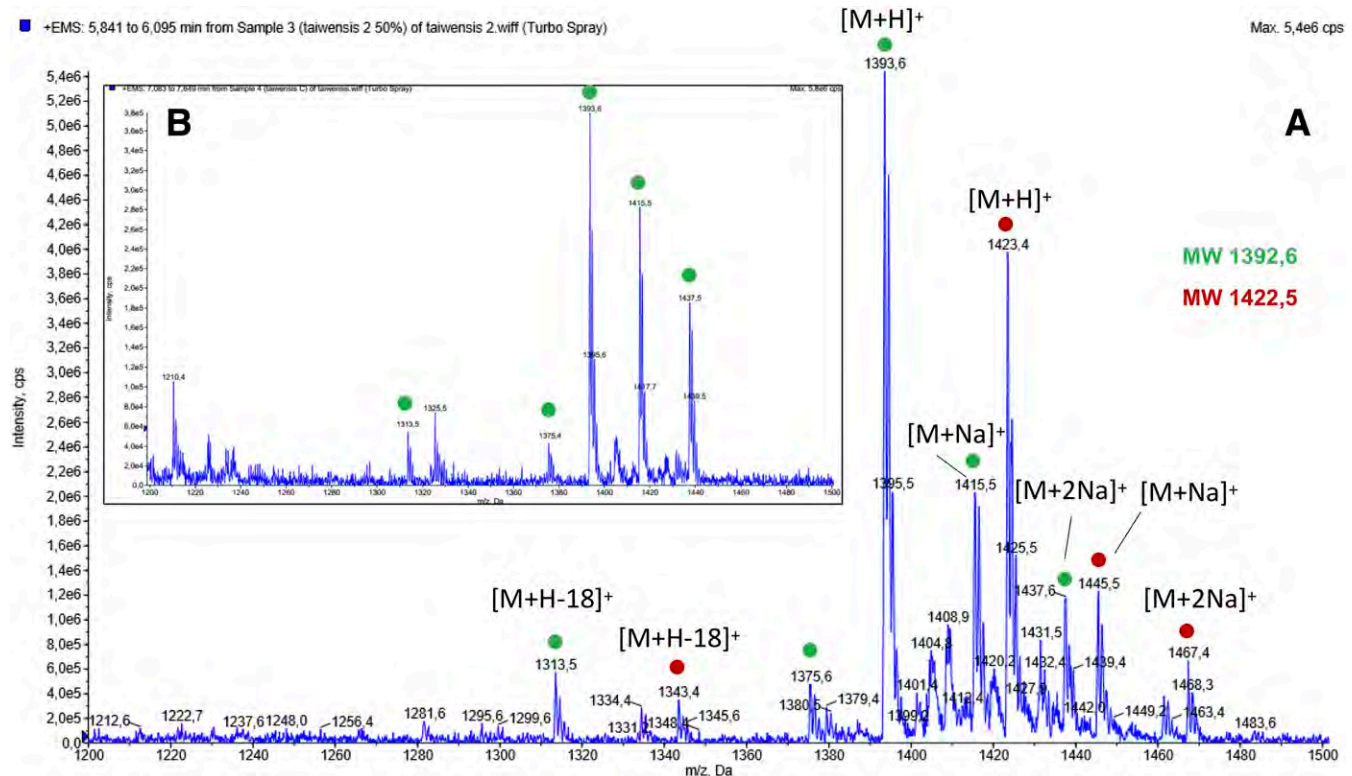


Fig. 4. Ultrahigh-performance liquid chromatography (UPLC) mass spectrometry spectra of *Cupriavidus taiwanensis* Nod factors (NFs). The reversed-phase SPE fraction eluting with 50% acetonitrile was subjected to tandem mass spectrometry analysis on a Q-Trap4500 system coupled with a UPLC device. The spectrum presented is the sum of the spectra obtained over the NF retention time domain. Even under strong acidic UPLC conditions, the presence of sodium adducts was dominant (e.g., MH⁺ at m/z 1,393 with MNa⁺ at m/z 1,415 and M(Na)₂⁺ at m/z 1,437 or 1,423.7 with 1,445.7 and 1,467.7, respectively). **A**, For the LMG19424 and its streptomycin-resistant derivative CBM832, two compound families were observed. **B**, For the deletion mutant CBM832 Δ *noeM*, only the usual NFs were detected.

Table 1. Summary of mass spectrometer analyses^a

Exp	Species	Strain	OD	Quantity LCO-V (counts/scan)	Quantity LCO redox (counts/scan)
1	<i>Cupriavidus taiwanensis</i>	LMG19424	0.881	2.92 e ⁶	8.50 e ⁶
1	<i>C. taiwanensis</i>	CBM832	0.678	6.50 e ⁶	3.99 e ⁶
2	<i>C. taiwanensis</i>	CBM832	0.8	3.25 e ⁶	2.70 e ⁶
3	<i>C. taiwanensis</i>	CBM832	0.71	2.00 e ⁶	6.80 e ⁶
4	<i>C. taiwanensis</i>	CBM832	0.918	26.00 e ⁶	18.00 e ⁶
5	<i>C. taiwanensis</i>	CBM832 Δ <i>noeM</i>	0.8	3.00 e ⁶	0
6	<i>C. taiwanensis</i>	CBM832 Δ <i>noeM</i>	0.27	0.57 e ⁶	0
7	<i>C. taiwanensis</i>	CBM832 Δ <i>noeM</i>	0.996	2.00 e ⁶	0
3	<i>C. taiwanensis</i>	CBM832 Δ <i>noeM</i>	0.668	12.60 e ⁶	0
4	<i>C. taiwanensis</i>	CBM832 Δ <i>noeM</i> :: <i>noeM</i>	0.842	21.40 e ⁶	12.40 e ⁶
4	<i>C. taiwanensis</i>	CBM832 Δ <i>noeM</i> :: <i>noeM</i>	0.991	22.50 e ⁶	15.00 e ⁶
8	<i>Rhizobium solanacearum</i>	GMI1000pRalta	0.443	17.2 e ⁶	8.40 e ⁶
9	<i>R. solanacearum</i>	GMI1000pRalta	0.599	5.30 e ⁶	10.20 e ⁶
9	<i>R. solanacearum</i>	GMI1000pRalta Δ <i>noeM</i>	0.52	0.74 e ⁶	0
9	<i>R. solanacearum</i>	GMI1000pRalta Δ <i>noeM</i>	0.546	0.93 e ⁶	0

^a Nod factors produced by *C. taiwanensis* and *R. solanacearum* derivatives. OD represents the optical cell density for each culture at 600 nm. LCO = lipochitoooligosaccharide. Quantities are given as ion intensities on the spectra summed over 1.1 min (Nod factor domain) and recorded after injecting 2 μ l of the SPE fraction dissolved in 250 μ l of acetonitrile/H₂O (1:1). Independent experiments (Exp) have different numbers.

S2). These proteins fall into two groups. The first group (group I) contained proteins exhibiting low similarity ($\leq 44\%$ identity in amino acid sequence, aligning $\geq 57\%$ of the query protein). Proteins of this group are mainly from non rhizobial strains belonging to the *Sphingomonas*, *Roseomonas*, *Devosia*, *Massilia*, or *Comamonas* genera but are also present in some rhizobial strains (*P. tuberum* WSM4176 and DUS833, *P. ribeironis* STM7296, and *R. freirei* PRF81). The second group (group II) contained proteins exhibiting strong similarity with pRalta_0471 ($\geq 70\%$ identity in amino acid sequence, aligning $\geq 76\%$ of the query protein). All of the proteins of this group are from rhizobial strains belonging to seven genera (i.e., *Ensifer*/*Sinorhizobium*, *Mesorhizobium*, *Neorhizobium*, and *Rhizobium* in α -proteobacteria and *Paraburkholderia*, *Trinickia*, and *Cupriavidus* in β -proteobacteria). Some strains in this group possess two copies (*Trinickia symbiotica* JPY366 and *Paraburkholderia* sp. CCGE1002 and JPY251). Group II proteins grouped into a strongly supported phylogenetic clade, clearly separated from other NoeM homologs (Fig. 5). Corresponding genes also form a separate phylogenetic clade (Supplementary Fig. S3). *C. taiwanensis* NoeM is predicted to be a transmembrane protein with three N-terminal transmembrane helices and to contain a C-terminal fatty acid hydroxylase domain of 134 amino acids (IPR006694/PF04116) with one HXXXH and two HXXHH motifs involved in iron binding (Shanklin et al. 1994) (Supplementary Fig. S4). These features are highly conserved in all proteins of group II, except in strains *R. freirei* PRF81, *R. tropici* CIAT899, *R. leucaenae* CFN299, *R. lusitanum* P1-7, and *Mesorhizobium helmanticense* CSLC115N, where the NoeM homologous protein corresponds to a 3' fragment of a full-length protein interrupted by a transposase, and is likely nonfunctional.

Among rhizobia possessing group II *noeM* homologs, 22 strains distributed in the *Rhizobium*, *Sinorhizobium*, *Mesorhizobium*, *Paraburkholderia*, and *Cupriavidus* genera have a complete assembled genome (Fig. 5). In all of these strains, the *noeM* gene is located on the symbiotic plasmid. The only exception is *M. huakuii* 7653R, where *noeM* is not located on the plasmid that bears the common *nodABC* genes but on another plasmid (pMHb), together with *nodM*. In all of these rhizobial strains, except in *R. tropici* CIAT899, where *noeM* is interrupted by a transposase, and in *Paraburkholderia* sp. CCGE1002 upstream from the second copy of the gene, *noeM* is preceded by a *nod* box. In these genomes, *noeM* is either clustered with some symbiotic genes (e.g., *R. etli* sv. *mimosae* Mim1) or away from them and framed by transposases (e.g., *C. taiwanensis* LMG19424), suggesting different evolutionary histories (Fig. 6).

Altogether, this showed that a class of fatty acid hydroxylases forming a separate phylogenetic clade are exclusively present in rhizobia and connected to the *nod* regulation. Thus, we called these genes and proteins that specialized in legume symbiosis *noeM* and NoeM, respectively.

***noeM* is overrepresented in *Mimosa* symbionts.**

Most of the rhizobial-retrieved NoeM sequences come from strains originally isolated from *Mimosa* spp. (48%) or from species belonging to the Mimoseae tribe (56%) of the Caesalpinioideae subfamily. To evaluate the link between *noeM* and *Mimosa* symbiosis, we screened our collection of *Mimosa* strains for the presence of *noeM* by PCR amplification and sequencing of an internal fragment of *noeM* (Table 2). We found that 22 of 35 (63%) of the *Mimosa* strains possess an *noeM* gene. Phylogenetic analysis confirmed that nucleotide sequences belong to the *noeM* clade. Notably, all ($n = 11$) of the *Mimosa pudica* strains harbored *noeM*, whatever their genus affiliation or geographical origin, suggesting that this gene is a feature of *M. pudica* symbionts. To explore this possibility, we

analyzed strains isolated from legumes belonging to other Mimoseae genera (*Anadenanthera*, *Piptadenia*, *Parapiptadenia*, *Prosopis*, and *Goldmania*) or to genera from other tribes (*Aspalathus*, *Dalbergia*, and *Phaseolus*) for both the presence of *noeM* and the capacity to nodulate *M. pudica* (Table 2). We found a strong correlation between these two traits. All strains nodulating *M. pudica* (i.e. 11 *Paraburkholderia* strains and 1 *Sinorhizobium arboris* strain isolated from legumes of the Mimoseae tribe) possessed *noeM*. Conversely, most strains possessing *noeM* nodulated *M. pudica*. Exceptions were strains *R. freirei* PFR81, *R. leucaenae* CFN299, *R. tropici* CIAT899, and *R. gallicum* sv. *gallicum* R602 isolated from *Phaseolus vulgaris*; *R. gallicum* UPRM8060 isolated from *Piptadenia flava*; and *S. arboris* LM14917 and *S. kostiense* LMG19225 isolated from *Prosopis chilensis*, which did not nodulate *M. pudica* in the laboratory. Notably, in *R. freirei* PFR81, *R. leucaenae* CFN299, and *R. tropici* CIAT899, *noeM* is interrupted by a transposase (see above).

The *noeM* gene of *C. taiwanensis* is involved in symbiosis with *M. pudica*.

To evaluate the role of *noeM* in the symbiotic capacity of *M. pudica* symbionts, we constructed a deletion mutant of *C. taiwanensis* CBM832, a streptomycin-resistant derivative of LMG19424.

M. pudica plantlets were single inoculated with CBM832 or its *noeM* mutant, and their nodulation followed over 28 days. No significant difference was observed between the nodulation kinetics of *C. taiwanensis* and of its *noeM* mutant (Supplementary Fig. S5). To more deeply analyze the role of *noeM* in the *C. taiwanensis*–*M. pudica* symbiosis, *M. pudica* plantlets were coinoculated with a mixture (1:1 ratio, 5×10^5 total bacteria per plant) of CBM832 constitutively expressing green fluorescent protein (GFP) and CBM832 Δ *noeM*. Nodules were collected at 21 and 42 days postinoculation (dpi). Each nodule was individually analyzed for the presence and the quantity of each strain by plating, allowing determination of the relative in planta fitness (percentage of each strain recovered from all nodules), the nodulation competitiveness (percentage of nodules inhabited by each strain), and the nodule infection capacity of each strain (number of bacteria per nodule for each strain). The *noeM* mutant was much less fit in planta than the wild-type strain, because mutant bacteria represented approximately 5.61 and 12.06% of the bacteria isolated from all nodules at 21 and 42 dpi, respectively (Fig. 7A). The differential fitness was due to a lower nodulation competitiveness of CBM832 Δ *noeM*. This strain formed only approximately 18.94% of total nodules at 21 dpi and 14.35% at 42 dpi (Fig. 7B). Yet, the CBM832 Δ *noeM* capacity to infect nodules it induces seemed not to be altered, because the numbers of bacteria per nodule were found to be similar for both strains at 21 and 42 dpi (Fig. 7C). Complementation of CBM832 Δ *noeM* with the *noeM* gene restored the in planta fitness and the nodulation competitiveness of the strain (Fig. 7A and B).

DISCUSSION

Bacterial *nod* genes play a key role in the establishment of nitrogen-fixing symbiosis by synthesizing the NF signaling compounds that trigger the plant developmental program of nodule organogenesis (Oldroyd and Downie 2008). Most NFs described to date are oligomers of N-acetyl-D-glucosaminyl residues, N-acylated on their nonreducing end, and substituted by various groups on their reducing or nonreducing residues (D'Haese and Holsters 2002). A very few atypical NFs that do not follow this rule have been described. As examples, 3-O fucosylation and 3-O acetylation of the glucosamine residue

next to the nonreducing residue were reported in *Mesorhizobium loti* (Olsthoorn et al. 1998) and *R. galegae* (Yang et al. 1999) NFs, respectively, and a glucose is inserted in the backbone of the molecule in *S. fredii* (Bec-Ferté et al. 1996). Here, we found that *C. taiwanensis* LMG19424 produces, in

addition to the compounds previously described (Amadou et al. 2008), novel NF compounds where the reducing sugar is in the open-chain form and oxidized (Fig. 3). This terminal residue carries two carboxyl, one N-acetyl, and one S groups. Yet, the arrangement of these chemical groups as proposed in Figure 3B

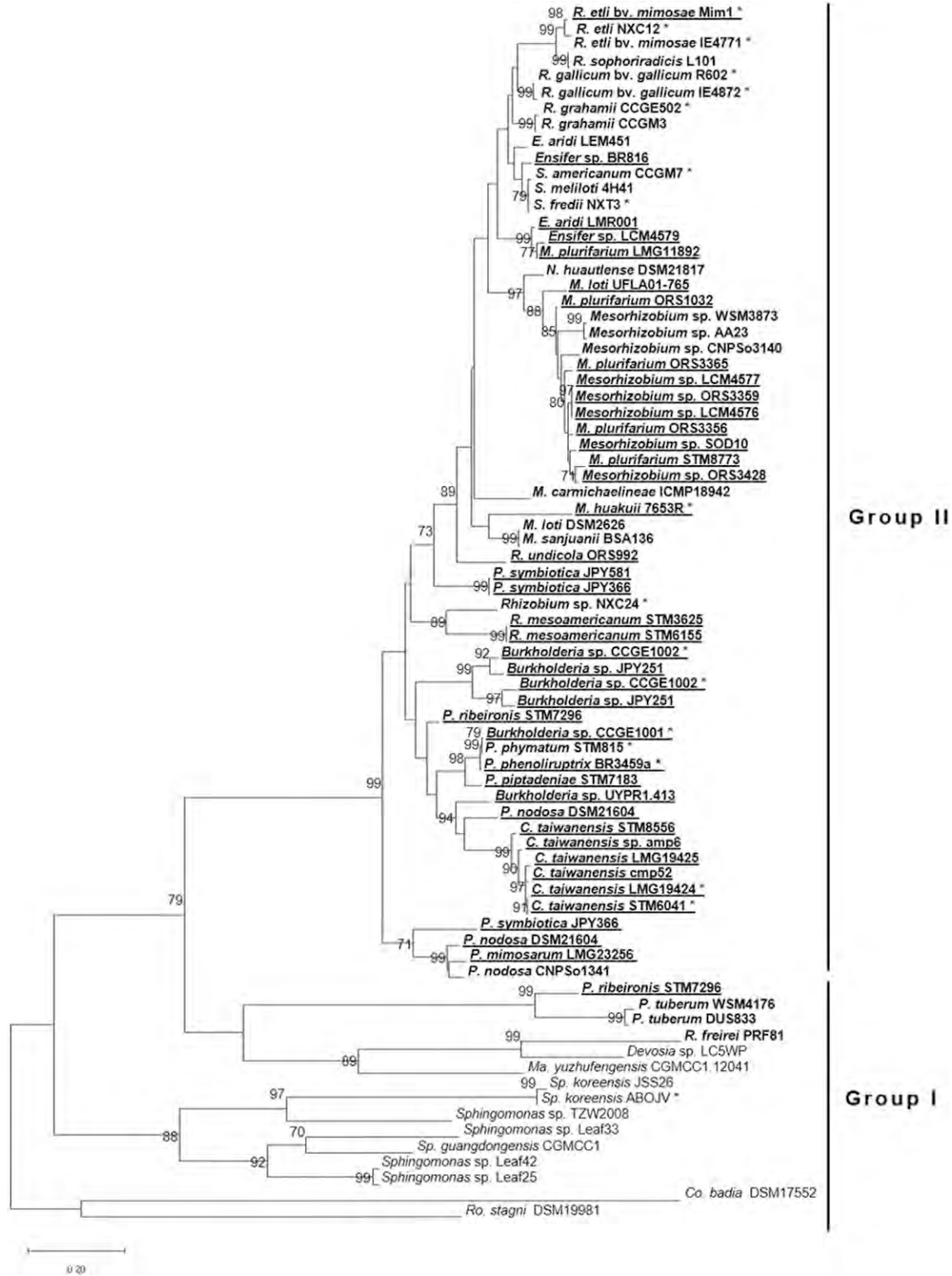


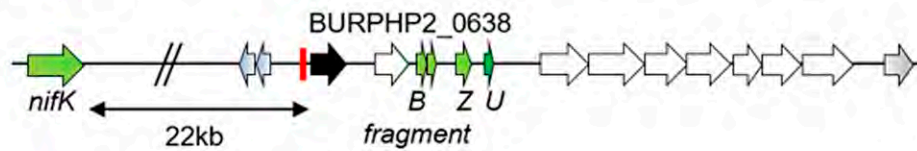
Fig. 5. Phylogenetic tree of protein sequences homologous to *Cupriavidus taiwanensis* NoeM. Maximum-likelihood phylogenetic tree based on 76 sequences with 626 amino acid positions inferred by the MEGA X program. Numbers at internal nodes correspond to bootstrap support values (in percentages) obtained in the analysis of 1,000 replicates using the same program and model of sequence evolution. Only bootstraps above 70 are indicated. Truncated sequences have been excluded from the analysis. Strains in bold are rhizobia, strains underlined were isolated from a plant belonging to the Mimoseae tribe, and an asterisk (*) indicates assembled genomes. C = *Cupriavidus*, Co = *Comamonas*, E = *Ensifer*, Ma = *Massilia*, M = *Mesorhizobium*, P = *Paraburkholderia*, R = *Rhizobium*, Ro = *Roseomonas*, S = *Sinorhizobium*, and Sp = *Sphingomonas*.

remains hypothetical, because complete assignment was not possible. The presence of an open-chain terminal residue was already described for the *Leucaena leucocephala* symbiont *Rhizobium* sp. BR816 (Supplementary Fig. S6) and the *Phaseolus vulgaris* symbiont *R. gallicum* sv. *gallicum* R602, where the “reducing” residue is a sulfated glucosaminitol with a varying degree of acetylation (Snoeck et al. 2001; Soria-Diaz et al. 2006). However, the terminal residue was found reduced in strains BR816 and R602, whereas it is oxidized with two carboxyl groups in *C. taiwanensis* LMG19424, increasing the complexity of the structure.

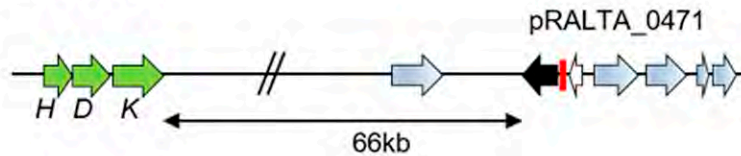
We identified a gene (*pRalta_0471*), which we called *noeM*, involved in the biosynthesis of these novel open-chain oxidized NF structures. A *C. taiwanensis* strain lacking this gene did not produce open-chain oxidized structures, whereas the reintroduction of *noeM* in this mutant restored the production of such structures. Furthermore, introduction of a wild-type or an *noeM*-deleted *pRalta* in *R. solanacearum* confirmed that *C. taiwanensis* NoeM is involved in the modification of the terminal residue. NoeM (275 amino acids) is predicted to belong to the fatty acid

hydroxylase superfamily (IPR006694/PF04116), a group of integral membrane proteins with an oxidoreductase activity and mainly involved in lipid biosynthesis. Enzymes of this superfamily, which includes fatty acid and carotene hydroxylases as well as sterol desaturases (Kim and Oh 2013), may thus include members acting on the sugar side of liposaccharides. *Rhizobium* sp. BR816 and *R. gallicum* R602, which both produce NFs with an open “reducing sugar”, also contain a very close NoeM homolog (76% amino acid identity, aligning 99% of the *C. taiwanensis* protein). However, this terminal residue is reduced, except if open-chain oxidized structures have been missed due to early technology. Thus, we hypothesize that *noeM* is involved in the production of an open-chain reduced terminal residue, while another still undetermined protein or proteins—present in *C. taiwanensis*—is involved in the further oxidation of this residue. Indeed, it is unlikely that a single enzyme catalyzes the two (several) steps of the LCO-V (C18:1/C16:0, NMe, Cb/H, and S) modification (i.e., aldehyde reduction of a spontaneously opened reducing sugar and the oxidation of the glucosaminitol group) after the migration of the sulfate to the O-5 position.

B. phymatum STM815



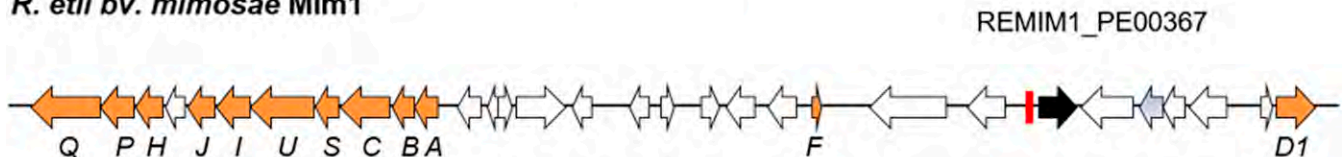
C. taiwanensis LMG 19424



R. mesoamericanum STM3625



R. etli bv. *mimosae* Mim1



R. grahamii CCGE502

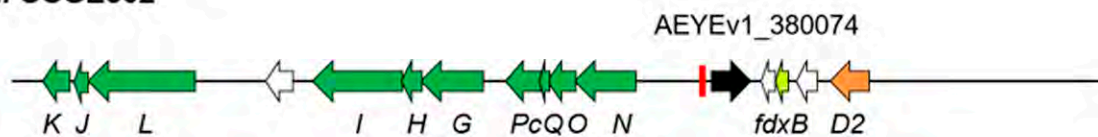


Fig. 6. Genomic context of *noeM* genes. Genes are represented by arrows of different colors and *nod* boxes are represented by red bars. Not all open reading frames are represented. Data for *Rhizobium grahamii* is from Althabegoiti et al. (2014).

Table 2. Presence of *noeM* and nodulation of *Mimosa pudica* (Nod) in a collection of legume isolates^a

Genus	Species	Strain	Original host plant	Host plant tribe	Geographic origin	Nod	<i>noeM</i>	Strain source/reference
<i>Bradyrhizobium</i>	sp.	STM7329	<i>Piptadenia paniculata</i>	Mimoseae	Brazil	–	0	Bournaud et al. 2013
<i>Bradyrhizobium</i>	sp.	STM7331	<i>Piptadenia paniculata</i>	Mimoseae	Brazil	–	0	Bournaud et al. 2013
<i>Bradyrhizobium</i>	sp.	STM7333	<i>Piptadenia paniculata</i>	Mimoseae	Brazil	–	0	Bournaud et al. 2013
<i>Bradyrhizobium</i>	sp.	STM7378	<i>Goldmania paraguensis</i>	Mimoseae	Brazil	–	0	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>caribensis</i>	STM1423	<i>Dalbergia louveli</i>	Dalbergieae	Madagascar	–	0	Rasolomampianina et al. 2005
<i>Paraburkholderia</i>	<i>caribensis</i>	STM1424	<i>Dalbergia louveli</i>	Dalbergieae	Madagascar	–	0	Rasolomampianina et al. 2005
<i>Paraburkholderia</i>	<i>caribensis</i>	TJ183	<i>Mimosa pudica</i>	Mimoseae	Taiwan	...	1	Chen et al. 2003a
<i>Paraburkholderia</i>	<i>diazotrophica</i>	STM7399	<i>Anadenanthera peregrina</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>diazotrophica</i>	STM7439	<i>Anadenanthera colubrina</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>mimosarum</i>	PAS44T	<i>Mimosa pigra</i>	Mimoseae	Taiwan	...	0	Chen et al. 2006
<i>Paraburkholderia</i>	<i>nodosa</i>	JPY258	<i>Mimosa albolanata</i>	Mimoseae	Brazil	...	1	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>nodosa</i>	JPY280	<i>Mimosa clausenii</i>	Mimoseae	Brazil	...	1	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>nodosa</i>	JPY629	<i>Mimosa cyclophylla</i>	Mimoseae	Brazil	...	1	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>nodosa</i>	STM7353	<i>Piptadenia trisperma</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>nodosa</i>	STM7363	<i>Parapiptadenia pterosperma</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>phenoliruptrix</i>	JPY580	<i>Mimosa cordistipula</i>	Mimoseae	Brazil	...	1	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>phenoliruptrix</i>	JPY582	<i>Mimosa hexandra</i>	Mimoseae	Brazil	...	1	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>phenoliruptrix</i>	STM7317	<i>Piptadenia gonocantha</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>phenoliruptrix</i>	STM7324	<i>Piptadenia paniculata</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>phenoliruptrix</i>	STM7415	<i>Anadenanthera peregrina</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>phenoliruptrix</i>	STM7454	<i>Anadenanthera colubrina</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>phymatum</i>	STM4206	<i>Mimosa pudica</i>	Mimoseae	French Guiana	...	1	Mishra et al. 2012
<i>Paraburkholderia</i>	<i>sabiae</i>	STM7319	<i>Piptadenia gonocantha</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>sabiae</i>	STM7419	<i>Anadenanthera peregrina</i>	Mimoseae	Brazil	+	1	Bournaud et al. 2013
<i>Paraburkholderia</i>	sp.	STM7296	<i>Piptadenia gonocantha</i>	Mimoseae	Brazil	+	1*	Bournaud et al. 2013
<i>Paraburkholderia</i>	<i>tuberum</i>	JPY303	<i>Mimosa velloziana</i>	Mimoseae	Brazil	...	0	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>tuberum</i>	JPY156	<i>Mimosa setosa</i>	Mimoseae	Brazil	...	0	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>tuberum</i>	JPY158	var. <i>paludosa</i>	Mimoseae	Brazil	...	0	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>tuberum</i>	JPY161	<i>Mimosa foliolosa</i>	Mimoseae	Brazil	...	0	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>tuberum</i>	JPY623	<i>Mimosa radula</i>	Mimoseae	Brazil	...	0	Bontemps et al. 2010
<i>Paraburkholderia</i>	<i>tuberum</i>	STM6020	<i>Mimosa pudica</i>	Mimoseae	French Guiana	...	1	Mishra et al. 2012
<i>Paraburkholderia</i>	<i>tuberum</i>	STM678	<i>Aspalathus carnosa</i>	Crotalariaeae	South Africa	–	0*	Moulin et al. 2001
<i>Cupriavidus</i>	sp.	MAPUD 10.1	<i>Mimosa pudica</i>	Mimoseae	Costa Rica	...	1	Andrus et al. 2012
<i>Cupriavidus</i>	sp.	UYMMa02A	<i>Mimosa magentea</i>	Mimoseae	Uruguay	...	1*	Iriarte et al. 2016
<i>Cupriavidus</i>	sp.	ip 2.35/pp2.35	<i>Mimosa diplotricha</i>	Mimoseae	Philippines	...	1	Andrus et al. 2012
<i>Cupriavidus</i>	sp.	MYM 10	<i>Mimosa diplotricha</i>	Mimoseae	Porto Rico	...	1	Andrus et al. 2012
<i>Cupriavidus</i>	sp.	pp 2.41	<i>Mimosa pudica</i>	Mimoseae	Philippines	...	1	Andrus et al. 2012
<i>Cupriavidus</i>	sp.	pp 2.75	<i>Mimosa pudica</i>	Mimoseae	Philippines	...	1	Andrus et al. 2012
<i>Cupriavidus</i>	sp.	SWF66194	<i>Mimosa pudica</i>	Mimoseae	China	...	1	Liu et al. 2011
<i>Cupriavidus</i>	<i>taiwanensis</i>	STM6077	<i>Mimosa pudica</i>	Mimoseae	New Caledonia	...	1	Klonowska et al. 2012
<i>Cupriavidus</i>	<i>taiwanensis</i>	STM6133	<i>Mimosa pudica</i>	Mimoseae	New Caledonia	...	1	Klonowska et al. 2012
<i>Ensifer</i>	sp.	MH37.1	<i>Mimosa himalayana</i>	Mimoseae	India	...	0	Gehlot et al. 2013
<i>Ensifer</i>	sp.	MHM12.2	<i>Mimosa hamata</i>	Mimoseae	India	...	0	Gehlot et al. 2013
<i>Mesorhizobium</i>	<i>chacoense</i>	LMG 19008T	<i>Prosopis alba</i>	Mimoseae	Argentina	–	0	Velazquez et al. 2001
<i>Mesorhizobium</i>	<i>chacoense</i>	LMG 19009	<i>Prosopis alba</i>	Mimoseae	Argentina	–	0	Velazquez et al. 2001
<i>Mesorhizobium</i>	<i>chacoense</i>	STM7859	<i>Prosopis alba</i>	Mimoseae	Argentina	–	0	Velazquez et al. 2001
<i>Rhizobium</i>	<i>etli</i>	TJ169	<i>Mimosa diplotricha</i>	Mimoseae	Taiwan	...	0	Chen et al. 2003a
<i>Rhizobium</i>	<i>etli</i> sv. <i>mimosae</i>	Mim-7-4	<i>Mimosa affinis</i>	Mimoseae	Mexico	...	1	Wang et al. 1999
<i>Rhizobium</i>	<i>freirei</i>	PRF81	<i>Phaseolus vulgaris</i>	Phaseolinae	Brazil	–	1*§	Ormeño-Orrillo et al. 2012
<i>Rhizobium</i>	<i>gallicum</i>	UPRM8060	<i>Piptadenia flava</i>	Mimoseae	Puerto Rico	–	1	Zurdo-Piñeiro et al. 2004
<i>Rhizobium</i>	<i>gallicum</i> sv. <i>gallicum</i>	R602T	<i>Phaseolus vulgaris</i>	Phaseolinae	France	–	1*	Geniaux et al. 1993
<i>Rhizobium</i>	<i>giardinii</i> sv. <i>giardini</i>	H152T	<i>Phaseolus vulgaris</i>	Phaseolinae	France	–	0	Amarger et al. 1997
<i>Rhizobium</i>	<i>leucaenae</i>	CFN299	<i>Phaseolus vulgaris</i>	Phaseolinae	Mexico	–	1*§	Ribeiro et al. 2012
<i>Rhizobium</i>	<i>mesoamericanum</i>	STM3629	<i>Mimosa pudica</i>	Mimoseae	French Guiana	...	1	Mishra et al. 2012
<i>Rhizobium</i>	<i>mesoamericanum</i>	STM6135	<i>Mimosa pudica</i>	Mimoseae	New Caledonia	...	1	Klonowska et al. 2012
<i>Rhizobium</i>	<i>mesoamericanum</i>	STM6161	<i>Mimosa pudica</i>	Mimoseae	New Caledonia	...	1	Klonowska et al. 2012
<i>Rhizobium</i>	sp.	STM7330	<i>Piptadenia paniculata</i>	Mimoseae	Brazil	–	0	Bournaud et al. 2013
<i>Rhizobium</i>	sp.	STM7379	<i>Mimosa foetida</i>	Mimoseae	Brazil	...	0	Bournaud et al. 2013
<i>Rhizobium</i>	sp.	STM7420	<i>Anadenanthera peregrina</i>	Mimoseae	Brazil	–	0	Bournaud et al. 2013
<i>Rhizobium</i>	sp.	STM7426	<i>Anadenanthera peregrina</i>	Mimoseae	Brazil	–	0	Bournaud et al. 2013
<i>Rhizobium</i>	sp.	TJ167	<i>Mimosa diplotricha</i>	Mimoseae	Taiwan	...	0	Chen et al. 2003a
<i>Rhizobium</i>	sp.	NGR181	<i>Mimosa invisa</i>	Mimoseae	Papua New Guinea	...	0	Elliott et al. 2009
<i>Rhizobium</i>	<i>tropicum</i>	CIAT 899	<i>Phaseolus vulgaris</i>	Phaseolinae	Colombia	–	1*§	Ormeño-Orrillo et al. 2012
<i>Rhizobium</i>	<i>tropicum</i>	TJ172	<i>Mimosa diplotricha</i>	Mimoseae	Taiwan	...	1	Chen et al. 2003a
<i>Rhizobium</i>	<i>tropicum</i>	UPRM8021	<i>Mimosa ceratonia</i>	Mimoseae	Puerto Rico	...	0	Zurdo-Piñeiro et al. 2004
<i>Sinorhizobium</i>	<i>arboris</i>	LMG 14917	<i>Prosopis chilensis</i>	Mimoseae	Kenya	–	1	Nick et al. 1999
<i>Sinorhizobium</i>	<i>arboris</i>	LMG 14919	<i>Prosopis chilensis</i>	Mimoseae	Soudan	+	1*	Nick et al. 1999
<i>Sinorhizobium</i>	<i>kostiense</i>	LMG 19225	<i>Prosopis chilensis</i>	Mimoseae	Soudan	–	1	Martens et al. 2008

^a The presence of *noeM* was analyzed by PCR, except for strains labeled with an asterisk (*), where the presence or absence of *noeM* was evidenced by genome sequence analysis. Symbol: § indicates that the *noeM* sequence corresponds to a truncated gene.

Further studies should confirm this hypothesis or assess which biochemical reaction is catalyzed by NoeM. NoeM, like NodC, is a transmembrane protein, while all other nodulation proteins involved in NFs synthesis are cytosolic. It has been proposed that the Nod proteins form a multienzyme complex anchored in the

inner membrane via NodC, allowing NF assembly on the cytoplasmic face of the inner membrane (Mergaert et al. 1997). NoeM could be part of the NodC complex or of another enzyme complex anchored in the inner membrane via this protein.

NoeM appeared to have specialized in rhizobial symbiosis. A phylogenetic analysis revealed that *C. taiwanensis* NoeM homologs split into two groups of proteobacterial proteins based on their homology with the query protein. The first group exhibited less than 44% identity in amino acid sequence while the second group exhibited more than 70%. Although homologs of the first group are distributed in rhizobial and non-rhizobial strains, homologs of the second group are only present in rhizobial species; specifically, species belonging to the genera *Ensifer/Sinorhizobium*, *Mesorhizobium*, *Neorhizobium*, and *Rhizobium* in α -proteobacteria and *Paraburkholderia*, *Trinickia*, and *Cupriavidus* in β -proteobacteria. In sequenced rhizobial genomes, *noeM* was found located in symbiotic plasmids downstream from a *nod* box, and sometimes clustered with other symbiotic (*nod* or *nif/fix/fixd*) genes. Moreover, in *C. taiwanensis*, *noeM* expression is induced by a *nodABC*-inducing flavonoid (luteolin) (Amadou et al. 2008) and is controlled by *nodD*. Notably, *noeM*, as *nodB* (Marchetti et al. 2011), is expressed all along the symbiosis, in bacteria colonizing the rhizosphere, infection sites, and infection threads as well as bacteroids. Expression in bacteroids was rarely observed in the rhizobium-legume symbiosis (Schlaman et al. 1991). Thus, we assigned the term *noeM* to these newly identified rhizobial genes that encode a protein involved in NF biosynthesis.

The specialization of a group of fatty acid hydroxylases into *nod* genes may have occurred in several steps (i.e., duplication of a gene of the fatty acid hydroxylase family, acquisition of biochemical activity on NFs if the original one did not possess it, and connection to the nodulation regulation pathway). *noeM* genes may then have spread to various rhizobial branches of α - and β -proteobacteria via horizontal gene transfer, resulting in incongruity between the taxonomical and *noeM* gene trees (Fig. 5). Most *nod* genes are monophyletic and are thought to have originated from such functional divergence and horizontal transfer (Aoki et al. 2013; Persson et al. 2015).

The *nod* genes encoding enzymes involved in the modification of the NF core structure have been shown to contribute nodulation efficiency or host specificity (Perret et al. 2000). Indeed, the presence of some host-specific *nod* genes has been correlated to the legume host spectrum (Lorquin et al. 1997). Although hundreds of rhizobial genomes have been sequenced (Joint Genome Institute Genomes OnLine Database), we only found *noeM* in 120 rhizobial genomes of the *Rhizobium*, *Sinorhizobium* (*Ensifer*), *Mesorhizobium*, *Neorhizobium*, *Paraburkholderia*, *Trinickia*, and *Cupriavidus* genera, 75% of them corresponding to isolates of the Mimoseae tribe. Specifically, *noeM* was identified in all *Mimosa pudica* isolates (this study) (Clerissi et al. 2018) and in all tested strains nodulating *M. pudica*, strongly suggesting that this gene plays a role in *M. pudica* symbiosis. Consistently, we found that it is required for optimal nodulation of *M. pudica* in *C. taiwanensis* because an *noeM*-deletion mutant had decreased nodulation competitiveness. Although *noeM* was not required for *M. pudica* nodulation in the laboratory, a role in nodulation efficiency may have a crucial impact in the field, where bacteria are in competition with many other rhizobia (Yates et al. 2011). However, the presence of a full-length *noeM* is not indicative of the capacity to nodulate *M. pudica*. This is the case of the *Phaseolus* symbiont *R. gallicum* sv. *gallicum* R602, which does not nodulate *M. pudica*, possesses *noeM*, and produces open-reduced NFs. The inability to nodulate *M. pudica* may be due to incompatibility factors such as secretion systems that prevent nodulation of some host plants (Nelson and Sadowsky 2015).

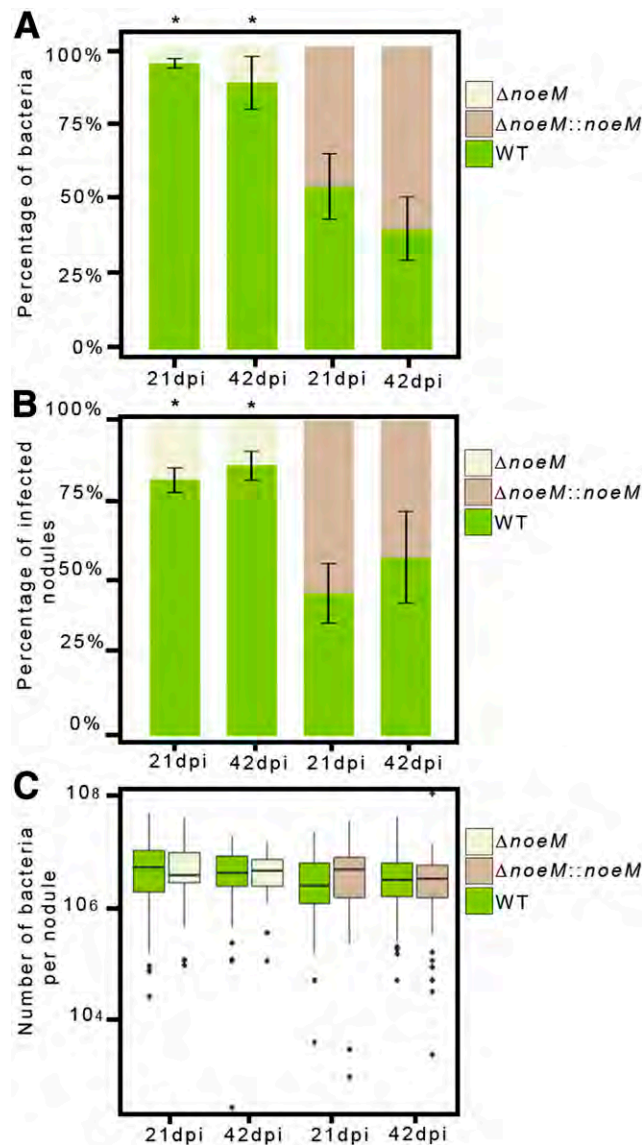


Fig. 7. Role of *noeM* in symbiosis with *Mimosa pudica*. *M. pudica* plantlets were coinoculated with a mix (1:1 ratio) of two *Cupriavidus taiwanensis* CBM832-derived strains, either wild-type green fluorescent protein (WT-GFP)/ $\Delta noeM$ or WT-GFP/ $\Delta noeM::noeM$. **A**, Relative in planta fitness; **B**, competitiveness for nodulation; and **C**, nodule infection capacity were analyzed by plating bacterial populations of nodules individually collected at 21 and 42 days postinoculation (dpi). In **A** and **B**, the percentage of bacteria or nodules is represented for each strain. Box plots in **C** represent the distribution of the number of bacteria recovered per nodule. Central rectangles span the first quartile to the third quartile (that is, the interquartile range), bold segments inside rectangles show the median, spots indicate suspected outliers, whiskers above and below the box show either the locations of the minimum and maximum in the absence of suspected outlying data. For the WT-GFP/ $\Delta noeM$ inoculum, 118 nodules were collected from 26 plants at 21 dpi and 136 nodules were collected from 23 plants at 42 dpi. For the WT-GFP/ $\Delta noeM::noeM$ inoculum, 121 nodules were collected from 23 plants at 21 dpi and 197 nodules were collected from 24 plants at 42 dpi. Data were from three independent experiments. An asterisk (*) indicates statistically significant difference between the two bacterial populations ($P < 0.05$, Student's *t* test with paired data at each time point for **A** and **B** and Wilcoxon rank sum test with continuity correction for **C**).

Future studies should determine whether the production of open-oxidized NFs is a feature of *M. pudica* symbionts.

MATERIALS AND METHODS

Bacterial strains and growth conditions.

Strains and plasmids used in this study are listed in Table 3. *C. taiwanensis* strains were grown at 28°C on tryptone-yeast (TY) medium supplemented with 6 mM CaCl₂ (Marchetti et al. 2010) or on minimal pseudo (MP) medium (Capela et al. 2017) supplemented with 10 mM succinate (MPS). *R. solanacearum* strains were grown at 28°C on B medium supplemented with 28 mM glucose (Capela et al. 2017) or on MP medium (Capela et al. 2017) supplemented with 10 mM glutamate. *E. coli* strains were grown at 37°C on Luria-Bertani medium. Antibiotics were used at the following concentrations: streptomycin at 200 µg/ml, trimethoprim at 100 µg/ml, tetracycline at 10 µg/ml, chloramphenicol at 12.5 µg/ml, carbenicillin at 50 µg/ml, and kanamycin at 25 µg/ml (for *E. coli*) or 50 µg/ml (for *R. solanacearum*).

Mutant construction.

Primers used for mutant constructions are listed in Supplementary Table S3. Unlabeled *noeM* and *nodD* deletion mutants of *C. taiwanensis* CBM832 were constructed using the

mutagenesis system developed by Flannagan et al. (2008), as described before (Daubech et al. 2017). Regions upstream and downstream from *noeM* were amplified with the oCBM3269-oCBM3271 and oCBM3272-oCBM3274 primer pairs using GoTaq DNA polymerase (Promega Corp.). The two PCR products were digested with *XbaI*-*Bam*HI and *Bam*HI-*Eco*RI, respectively, and cloned into the pGPI-*Sce*I plasmid digested by *Xba*I and *Eco*RI. Ligation products were transformed into a DH5α λpir *E. coli* strain. The resulting plasmid was transferred into CBM832 by triparental mating using pRK600 as helper plasmid. Deletion of *nodD* was performed in the same way, using the oCBM3365-oCBM3366 and oCBM3367-oCBM3368 primer pairs. Transconjugants that have integrated the plasmid by single crossing over were selected on streptomycin and trimethoprim and verified by PCR using the oCBM3315-oCBM3274 and oCBM3316-oCBM3269 primer pairs (*noeM*) or the primers oCBM3358-oCBM3368 and oCBM3363-oCBM3365 (*nodD*). The pDAI-*Sce*I replicative plasmid was then introduced into these strains by conjugation using pRK2013 as helper plasmid and selected on tetracyclin. Expression of the I-*Sce*I nuclease causes a double strand break into the inserted plasmid and promotes DNA recombination. Deleted mutants were screened for trimethoprim sensitivity and verified by PCR using the oCBM3315/oCBM3316 (*noeM*) or oCBM3358/oCBM3363 (*nodD*) primer pairs. Mutants

Table 3. Strains and plasmids used in this study

Bacterium, plasmid	Relevant characteristics ^a	Reference or source
<i>Escherichia coli</i>		
DH5α	<i>F recA lacZ</i> M15, Car ^R	Bethesda Research Laboratory
DH5α λpir	<i>F recA lacZ</i> M15 λpir, Tri ^R	HP Schweizer
<i>Cupriavidus taiwanensis</i>		
CBM832	LMG19424 derivative, Str ^R	Marchetti et al. 2010
CBM2700	CBM832 carrying a <i>psbA</i> promoter (<i>Pps</i>)-green fluorescent protein (GFP) fusion downstream <i>glmS</i> , Str ^R	Daubech et al. 2017
CBM2708	CBM832 carrying a <i>Pps</i> -mCherry fusion downstream <i>glmS</i> , Str ^R	Daubech et al. 2017
CBM2720	CBM832 carrying the <i>nifH</i> promoter region (<i>PnifH</i>) fused to mCherry downstream <i>glmS</i> , Str ^R	This study
CBM2856	CBM832 carrying a 549-bp internal deletion in <i>noeM</i> , Str ^R	This study
CBM2858	CBM832 carrying a 404-bp DNA fragment upstream from <i>noeM</i> (<i>PnoeM</i>) fused to mCherry downstream from <i>glmS</i> , Str ^R	This study
CBM2927	CBM832 carrying a 642-bp internal deletion in <i>nodD</i> , Str ^R	This study
CBM2964	CBM2856 carrying a 234-bp DNA fragment including <i>noeM</i> and its promoter region downstream from <i>glmS</i> , Str ^R	This study
CBM2988	CBM2927 carrying a 404-bp DNA fragment <i>PnoeM</i> fused to mCherry downstream from <i>glmS</i> , Str ^R	This study
<i>Rhizobium solanacearum</i>		
CBM124	GMI1000 pRalta::Tri, Tri ^R	Marchetti et al. 2010
CBM2689	CBM124 derivative carrying a 549-bp internal disruption by a kanamycin cassette in <i>noeM</i> , Tri ^R , Kan ^R	This study
Plasmids		
pGEM-T	Cloning vector	Promega Corp.
pGPI- <i>Sce</i> I	<i>ori</i> _{R6K} , mob ⁺ , carries an I- <i>Sce</i> I site, Tri ^R	Flannagan et al. 2008
pDAI- <i>Sce</i> I	<i>ori</i> _{pBBR1} , mob ⁺ , carries the I- <i>Sce</i> I gene, Tet ^R	Flannagan et al. 2008
pCM162	pGPI- <i>Sce</i> I carrying the <i>glmS</i> -Ralta_A0206 intergenic region interrupted by a <i>Pps</i> -mCherry fusion, Tri ^R	Daubech et al. 2017
pCM164	pGPI- <i>Sce</i> I carrying the <i>glmS</i> -Ralta_A0206 intergenic region interrupted by an <i>nifH</i> -mCherry fusion, Tri ^R	This study
pCBM197	pGPI- <i>Sce</i> I carrying the <i>glmS</i> -Ralta_A0206 intergenic region interrupted by an <i>noeM</i> (<i>C. taiwanensis</i>)-mCherry fusion, Tri ^R	This study
pCBM198	pGPI- <i>Sce</i> I carrying the <i>C. taiwanensis noeM</i> 5' and 3' regions, Tri ^R	This study
pCBM199	pGPI- <i>Sce</i> I carrying the <i>nodD</i> 5' and 3' regions, Tri ^R	This study
pCBM201	pGPI- <i>Sce</i> I carrying the <i>glmS</i> -Ralta_A0206 intergenic region interrupted by <i>noeM</i> from <i>C. taiwanensis</i> and its promoter region, Tri ^R	This study
pCBM202	pGEM-T carrying carrying <i>noeM</i> from <i>C. taiwanensis</i> and a 404-bp DNA fragment upstream from the gene	This study
pRK600	Helper plasmid, Chl ^R	Finan et al. 1986
pRK2013	Helper plasmid, Kan ^R	Figurski and Helinski 1979
pJQ200	<i>ori</i> _{p15A} , mob ⁺ (RP4 <i>oriT</i>), Gm ^R , suicide plasmid	Quandt and Hynes 1993

^a Car^R, Tri^R, Str^R, Kan^R, Tet^R Chl^R, and Gm^R indicate resistant to carbenicillin, trimethoprim, streptomycin, kanamycin, tetracycline, chloramphenicol, and gentamicin, respectively.

were cultivated on unselective TY medium and tetracycline-sensitive colonies having lost the pDAI-*SceI* plasmid were selected.

An unlabeled *C. taiwanensis noeM* mutant complemented with *noeM* was constructed by inserting a 1,234-bp DNA fragment including *noeM* and its 404-bp upstream region in the CBM2856 intergenic region between the *glmS* and *RALTA_A0206* genes using the same pGPI-*SceI*/pDAI-*SceI* mutagenesis system described above. Regions upstream and downstream from *noeM* were amplified with oCBM3279-oCBM3430 using the phusion DNA polymerase (Thermo Fisher Scientific). The PCR products were digested with *AvrII* and *SpeI*, respectively, and cloned into pCM162 digested by *AvrII* and *SpeI*. Ligation products were transformed into a DH5 α λ pir *E. coli* strain and then transferred into *C. taiwanensis* CBM2856 by triparental mating using pRK600 as helper plasmid. Transconjugants were verified by PCR using the oCBM2739-oCBM3430 and oCBM3279-oCBM2740 primer pairs. Integration of the DNA fragment by crossing over was carried out using the pDAI-*SceI* plasmid as described above and verified by PCR using the oCBM2739-oCBM2740 primer pair.

The *nifH*-mCherry and *noeM*-mCherry fusions were inserted into the wild type and into the *nodD* mutant of *C. taiwanensis* at the same chromosomal locus (i.e., in the intergenic region between the *glmS* and *RALTA_A0206* genes using the same pGPI-*SceI*/pDAI-*SceI* mutagenesis system. The fusions were constructed using pCBM162 digested by *EcoRI* and *SpeI*. The promoter regions of *noeM* and *nifH* were amplified by PCR using the oCBM3279-oCBM3165 and the oCBM2617-oCBM2618 primer pairs, respectively. The PCR products were digested by *EcoRI* and *SpeI* and cloned in the digested pGPI, then inserted into the strains. Transconjugants were verified by PCR using the oCBM2739-oCBM2618/oCBM2617-oCBM2740 (*nifH*-mCherry fusions) or the oCBM2739-oCBM3165/oCBM3279-oCBM2740 (*noeM*-mCherry fusions) primer pairs. Integration of the DNA fragment by crossing over was carried out using the pDAI-*SceI* plasmid as described above and verified by PCR using the oCBM2739-oCBM2740 primer pair.

For the construction of the *noeM* deletion mutant of *R. solanacearum* CBM124, the 5' and 3' flanking regions were PCR amplified using the oCBM3388-oCBM3370 and oCBM3371-oCBM3372 primer pairs, respectively. The PCR products were digested with *EcoRI*/*Acc65I* and *SacI*/*SacII*, respectively, and cloned one after the other into pCM184 (Marx et al. 2002) digested with *EcoRI*/*Acc65I* and *SacI*/*SacII*. Ligation products were transformed into a DH5 α λ *E. coli* strain. The plasmid was linearized with *EcoRI* and introduced into *R. solanacearum* CBM124 by natural transformation, as described by Capela et al. (2017). The transformants were selected on kanamycin and the construction was verified using the oCBM3315-oCBM3316 primer pair.

NF purification and characterization.

Strains were grown in minimal media overnight, then diluted to OD = 0.1 and cultivated for 5 h, then diluted to OD = 0.05 in 1 liter of MPS and cultivated 2 h before adding 10 μ M luteolin. Cultures were grown for an additional 24 h and stopped at midexponential phase. Cultures were centrifuged at 3,500 rpm and supernatants extracted twice with 1-butanol. The organic fractions were pooled, washed with water, and vacuum dried. After dissolution in water and extraction with ethyl acetate, the fraction was dried again. The obtained residue was submitted to an SPE cleaning on a C18 5- μ m reversed-phase stationary phase (Macherey and Nagel) conditioned with water. Three fractions were collected: the 20% acetonitrile (ACN) in water, 50% ACN in water, and finally 100% ACN elution. The fractions were then submitted to mass spectrometric analyses. The mass spectrometer was a QTrap4500 (AB Sciex) equipped with

a UPLC device (Dionex Thermo Fisher Scientific). Analyses were performed in the less-selective positive mode. To increase the signal, enhanced mass spectra were acquired using the linear trap. MS/MS spectra were also acquired to assess the structure of the different compounds.

Plant tests.

M. pudica seed (B&T World Seed) were sterilized as described (Chen et al. 2003b). Seedlings were cultivated in Gibson tubes, as previously described (Marchetti et al. 2014). Plantlets were inoculated with a single strain for nodulation kinetics (0.5×10^6 bacteria/plant) or with a mixture of *C. taiwanensis* derivatives (GFP-labeled and unlabeled Δ *noeM* mutant, ratio 1:1; 10^6 total bacteria/tube) for in planta fitness, nodulation competitiveness, and nodule infection capacity analysis. Nodules of coinoculated plants were individually collected at 21 or 42 dpi, surface sterilized for 15 min in a 2.5% sodium hypochlorite solution, rinsed with water, and crushed. Each nodule crush was diluted in water and plated using an easy spiral automatic plater (Interscience). Green colonies (GFP-labeled strains) were counted after 2 days of incubation at 28°C, under a fluorescence stereo zoom microscope (Axiozoom V16; Zeiss) when appropriate. From these data, we calculated the in planta fitness (total bacteria recovered divided by the number of total nodules), the nodulation competitiveness (number of nodules inhabited by a given strain divided by the number of total nodules), and the nodule infection capacity (number of bacterial cells per nodule inhabited by a given strain) for each inoculated strain.

To analyze the expression of *noeM* during symbiosis, plants were grown on 12-cm² plates (three plants per plate) containing slanting nitrogen-free Fahraeus agar medium for 3 days at 28°C. Roots were covered with a sterile, gas-permeable, and transparent plastic film (BioFolie 25; Sartorius AG, Vivascience). Plants were inoculated with 10^7 mCherry-tagged bacteria per plant. Roots were collected at 24 h post-inoculation and 7 dpi and fixed in 3.2% paraformaldehyde. The fixed roots were observed using a confocal microscope (Confocal Leica SP8). Nodules were collected at 21 dpi and cut for nodule sections 55 mm thick using a Leica VT1000S vibratome. Samples were observed by using a fluorescence microscope (Zeiss AxioPhot).

Quantitative RT-PCR.

Bacteria were cultivated overnight in 30 ml of MPS medium. Overnight cultures were diluted to OD = 0.1 in 100 ml of MPS and cultivated for additional 2 h at 28°C. Cultures were then divided into two equal cultures of 50 ml, and luteolin (10 μ M final concentration) or absolute ethanol was added. Cultures were harvested 5 h 30 min after luteolin induction by adding 1/10 volume of cold phenol/ethanol (5:95) mixture, then centrifuged for 15 min at 4°C at 4,000 rpm. The pellets were frozen in liquid nitrogen and stored at -80°C until RNA extraction. For RNA extraction, pellets were resuspended in lysis-enzyme at 1 mg/ml in Tris-EDTA (10:1, pH 8) buffer and incubated for 5 min at room temperature. A mirVana kit (Ambion) was used for the extraction of the RNAs. To remove residual genomic DNA, the RNAs were treated at 37°C for 30 min with Turbo DNase and purified with the Sureprep kit (Fisher BioReagents) twice. The quality of the extracted RNAs was assessed using a 2100 Bioanalyzer instrument (Agilent Technologies) and their concentration was determined using a Nanodrop ND 1000 spectrophotometer (ThermoFisher Scientific). For quantitative RT-PCR, 1 μ g of RNA was reverse transcribed using the Transcriptor Reverse Transcription kit (Roche, Life Technologies) and random hexamers as primers (Invitrogen). Real-time PCRs were run on a LightCycler system

(Roche, Life Technologies) using the Eurogentec Takyon qPCR kit (Eurogentec). The *hisB* and *rpoA* genes were used as reference genes. Primer pairs oCBM3812/foCBM3813 for *hisB*, oCBM3814/oCBM3815 for *rpoA*, oCBM1513/oCBM1514 for *nodB*, and oCBM1515/oCBM1516 for *noeM* were used for quantitative PCR.

Quantification of the mCherry expression.

Bacteria were cultivated overnight in MPS, diluted to OD = 0.05, and cultivated for an additional 3 h. Luteolin (10 μ M final concentration) or the same volume of ethanol absolute was added and 200 μ l from each culture was distributed in triplicate on 96-well microplates. Fluorescence and growth were measured using The FLUOstar Omega (BMG LABTECH) over 24 h.

noeM homologs, symbiotic plasmids, and *nod* boxes.

We searched for pRalta_0471 homologs in Refseq proteins using blastP (searched on 10 October 2018). We discarded hits with e-values < 10⁻⁵, with <40% identity in amino acid sequence, or aligning <50% of the query protein. Conversely, homologous proteins were blasted against the *C. taiwanensis* proteome to identify bidirectional best hits (BBH). BBH lacking stops and frameshifts and aligning >70% of the query protein were defined as putative orthologs. Hits exhibiting >70% identity in amino acid sequence and aligning >70% of the query protein were defined as NoeM proteins. Plasmids were identified as symbiotic plasmids when they bore the common *nodABC* genes. For *nod* box prediction, 1 kb upstream from *noeM* were extracted from genome sequences and the AT(N11)AT(N7)AT(N11)AT motif was searched using Patscan (Dsouza et al. 1997) and confirmed by expert examination. For the screening of the presence of *noeM* homologs in rhizobial strains, DNA was extracted using a quick Proteinase K method (Wilson 1987) and amplified using the primer pairs St_desF1 and St_desR2, which allowed the most efficient universal amplification of an internal *noeM* 650-bp sequence. PCR amplifications were performed with GoTaq DNA Polymerase (Promega Corp.) according to the manufacturer's instructions, and using 0.625 U of GoTaq, 2.5 μ mol each dNTP, and 20 μ mol each primer in 25 μ l of volume reaction. PCR protocol included 1 min of initial DNA denaturation at 95°C; 35 cycles with 30 s of denaturation at 95°C, 30 s of hybridization at 55°C, and 1 min of elongation at 72°C; and a final step of 10 min at 72°C. PCR amplification products were checked for purity by electrophoresis, and sent for sequencing with the St_desR2 primer.

Alignment and phylogenetic analyses.

The phylogenetic tree of NoeM protein sequences was built using the the MEGA X program (Kumar et al. 2018). Multiple alignment was performed using Muscle. The evolutionary history of sequences was inferred by using the maximum-likelihood method (Le and Gascuel 2008). The tree with the highest log likelihood (-13,534.96) is shown. Initial trees for the heuristic search were obtained automatically by applying neighbor-joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, then selecting the topology with superior log-likelihood value. The multiple sequence alignment of rhizobial NoeM was performed using Multalin, version 5.4.1.

Statistical analyses.

For the statistical analysis of the mCherry expression, growth curves were fitted to logistic equations using the R package GrowthCurver (version 0.3.0) to obtain interpretable metrics from microbial growth curves (Sprouffske and Wagner 2016). The area under the logistic curve (AUC), which provides a

metric (the logistic AUC) that integrates information from the logistic parameters, was used in a two-way analysis of variance (ANOVA) examining the effect of the luteolin on *psbA::mCherry* and *noeM::mCherry*, taking into account biological repeats. Assumptions of the ANOVA were verified ($\alpha = 0.05$) using a Shapiro test for error normality distribution and a Bartlett test for error homoscedasticity. Post-hoc analyses were performed using the Tukey honestly significant difference test ($\alpha = 0.05$).

For the statistical analysis of the nodulation kinetics, the area under the curves of each experiment was calculated using the the R package agricolae (version 1.3-1) and compared between strains using Student's test in R.

ACKNOWLEDGMENTS

We thank M. Parker, P. Young, E. James, and E. Martínez-Romero for providing rhizobial strains, P. Remigi and M.-F. Jardinaud for help in statistical analyses, and S. Fort, S. Cottaz, and A. Rousseau for helpful discussion.

AUTHOR-RECOMMENDED INTERNET RESOURCE

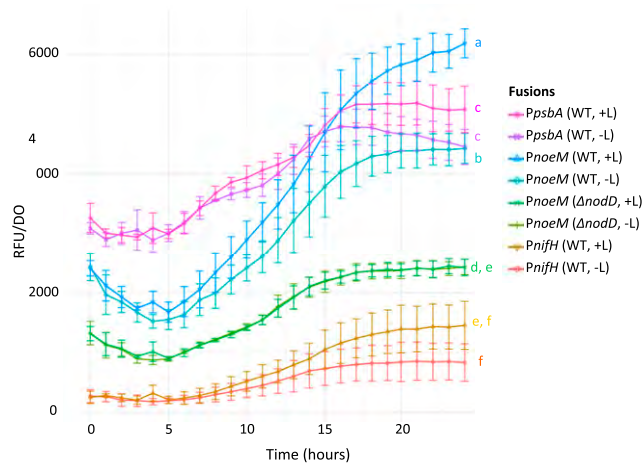
Joint Genome Institute Genomes OnLine Database: <https://gold.jgi.doe.gov/>

LITERATURE CITED

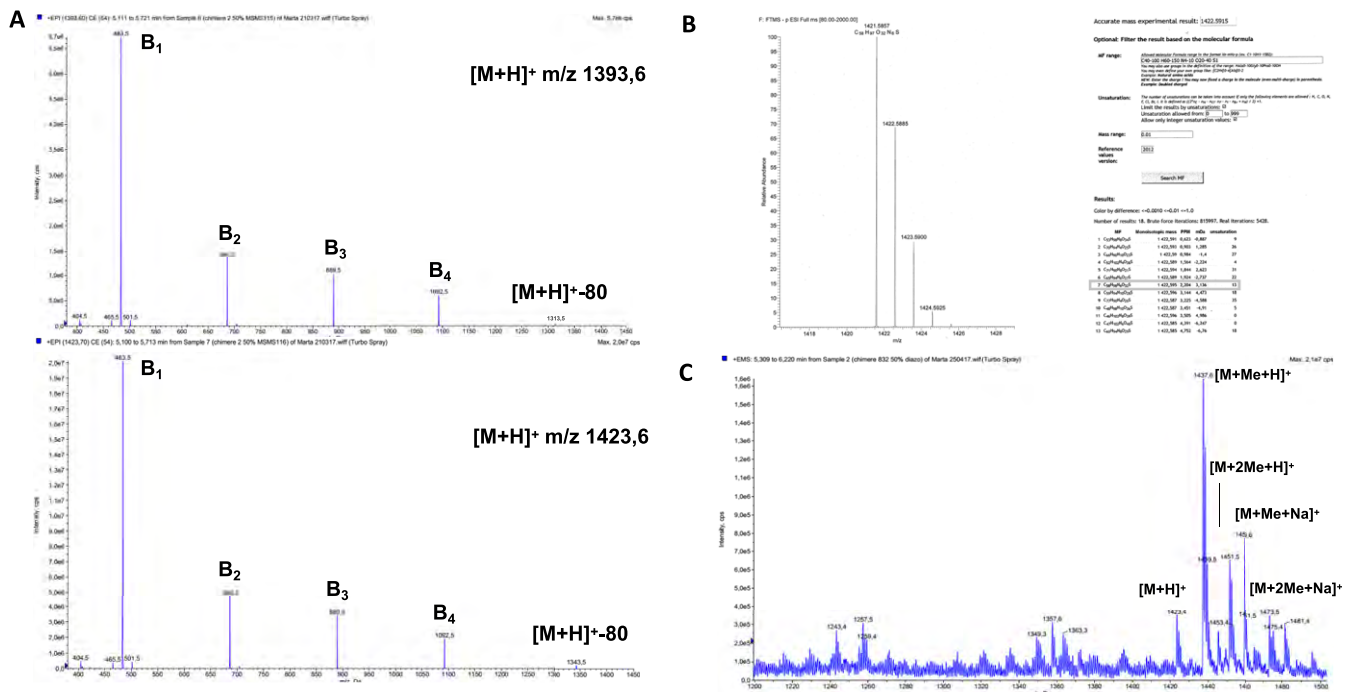
- Althabegoiti, M. J., Ormeño-Orrillo, E., Lozano, L., Torres Tejerizo, G., Rogel, M. A., Mora, J., and Martínez-Romero, E. 2014. Characterization of *Rhizobium grahamii* extrachromosomal replicons and their transfer among rhizobia. *BMC Microbiol.* 14: Article 6.
- Amadou, C., Pascal, G., Mangenot, S., Glew, M., Bontemps, C., Capela, D., Carrère, S., Cruveiller, S., Dossat, C., Lajus, A., Marchetti, M., Poinot, V., Rouy, Z., Servin, B., Saad, M., Schenowitz, C., Barbe, V., Batut, J., Médigue, C., and Masson-Boivin, C. 2008. Genome sequence of the β -rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Res.* 18:1472-1483.
- Amarger, N., Macheret, V., and Laguerre, G. 1997. *Rhizobium gallicum* sp. nov. and *Rhizobium giardinii* sp. nov., from *Phaseolus vulgaris* nodules. *Int. J. Syst. Bacteriol.* 47:996-1006.
- Andrus, A. D., Andam, C., and Parker, M. A. 2012. American origin of *Cupriavidus* bacteria associated with invasive *Mimosa* legumes in the Philippines. *FEMS Microbiol. Ecol.* 80:747-750.
- Aoki, S., Ito, M., and Iwasaki, W. 2013. From β - to α -proteobacteria: The origin and evolution of rhizobial nodulation genes *nodII*. *Mol. Biol. Evol.* 30:2494-2508.
- Bec-Ferté, M.-P., Krishnan, H. B., Savagnac, A., Pueppke, S. G., and Promé, J.-C. 1996. *Rhizobium fredii* synthesizes an array of lipooligosaccharides, including a novel compound with glucose inserted into the backbone of the molecule. *FEBS Lett.* 393:273-279.
- Bontemps, C., Elliott, G. N., Simon, M. F., Dos Reis Júnior, F. B., Gross, E., Lawton, R. C., Neto, N. E., de Fátima Loureiro, M., De Faria, S. M., Sprent, J. I., James, E. K., and Young, J. P. W. 2010. *Burkholderia* species are ancient symbionts of legumes. *Mol. Ecol.* 19:44-52.
- Bournaud, C., de Faria, S. M., dos Santos, J. M., Tisseyre, P., Silva, M., Chaintreuil, C., Gross, E., James, E. K., Prin, Y., and Moulin, L. 2013. *Burkholderia* species are the most common and preferred nodulating symbionts of the *Piptadenia* group (tribe Mimoseae). *PLoS One* 8:e63478.
- Brogghammer, A., Krusell, L., Blaise, M., Sauer, J., Sullivan, J. T., Maolanon, N., Vinther, M., Lorentzen, A., Madsen, E. B., Jensen, K. J., Roepstorff, P., Thirup, S., Ronson, C. W., Thygesen, M. B., and Stougaard, J. 2012. Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding. *Proc. Natl. Acad. Sci. U.S.A.* 109:13859-13864.
- Capela, D., Marchetti, M., Clérissi, C., Perrier, A., Guetta, D., Gris, C., Valls, M., Jauneau, A., Cruveiller, S., Rocha, E. P. C., and Masson-Boivin, C. 2017. Recruitment of a lineage-specific virulence regulatory pathway promotes intracellular infection by a plant pathogen experimentally evolved into a legume symbiont. *Mol. Biol. Evol.* 34:2503-2521.
- Chen, W. M., James, E. K., Coenye, T., Chou, J. H., Barrios, E., de Faria, S. M., Elliott, G. N., Sheu, S. Y., Sprent, J. I., and Vandamme, P. 2006. *Burkholderia mimosarum* sp. nov., isolated from root nodules of *Mimosa* spp. from Taiwan and South America. *Int. J. Syst. Evol. Microbiol.* 56: 1847-1851.

- Chen, W. M., James, E. K., Prescott, A. R., Kierans, M., and Sprent, J. I. 2003a. Nodulation of *Mimosa* spp. by the β -proteobacterium *Ralstonia taiwanensis*. *Mol. Plant-Microbe Interact.* 16:1051-1061.
- Chen, W. M., Moulin, L., Bontemps, C., Vandamme, P., Béna, G., and Boivin-Masson, C. 2003b. Legume symbiotic nitrogen fixation by β -proteobacteria is widespread in nature. *J. Bacteriol.* 185:7266-7272.
- Chen, X. C., Feng, J., Hou, B. H., Li, F. Q., Li, Q., and Hong, G. F. 2005. Modulating DNA bending affects NodD-mediated transcriptional control in *Rhizobium leguminosarum*. *Nucleic Acids Res.* 33:2540-2548.
- Clerissi, C., Touchon, M., Capela, D., Tang, M. X., Cruveiller, S., Parker, M. A., Moulin, L., Masson-Boivin, C., and Rocha, E. P. C. 2018. Parallels between experimental and natural evolution of legume symbionts. *Nat. Commun.* 9: Article 2264.
- Daubech, B., Remigi, P., Doin de Moura, G., Marchetti, M., Pouzet, C., Ariaci, M. C., Gokhale, C. S., Masson-Boivin, C., and Capela, D. 2017. Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation. *eLife* 6:e28683.
- D'Haese, W., and Holsters, M. 2002. Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology* 12: 79R-105R.
- Downie, J. A. 1998. Functions of rhizobial nodulation genes. Pages 387-402 in: *The Rhizobiaceae*. H. P. Spaink, A. Kondorosi, and P. J. J. Hooykaas, eds. Springer, Dordrecht, The Netherlands.
- Dsouza, M., Larsen, N., and Overbeek, R. 1997. Searching for patterns in genomic data. *Trends Genet.* 13:497-498.
- Elliott, G. N., Chou, J. H., Chen, W. M., Bloembergen, G. V., Bontemps, C., Martínez-Romero, E., Velázquez, E., Young, J. P. W., Sprent, J. I., and James, E. K. 2009. *Burkholderia* spp. are the most competitive symbionts of *Mimosa*, particularly under N-limited conditions. *Environ. Microbiol.* 11:762-778.
- Figurski, D. H., and Helinski, D. R. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc. Natl. Acad. Sci. U.S.A.* 76:1648-1652.
- Finan, T. M., Kunkel, B., De Vos, G. F., and Signer, E. R. 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J. Bacteriol.* 167:66-72.
- Flannagan, R. S., Linn, T., and Valvano, M. A. 2008. A system for the construction of targeted unmarked gene deletions in the genus *Burkholderia*. *Environ. Microbiol.* 10:1652-1660.
- Gehlert, H. S., Tak, N., Kaushik, M., Mitra, S., Chen, W. M., Poweleit, N., Panwar, D., Poonar, N., Parihar, R., Tak, A., Sankhla, I. S., Ojha, A., Rao, S. R., Simon, M. F., dos Reis Junior, F. B., Perigolo, N., Tripathi, A. K., Sprent, J. I., Young, J. P., James, E. K., and Gyaneshwar, P. 2013. An invasive *Mimosa* in India does not adopt the symbionts of its native relatives. *Ann. Bot.* 112:179-196.
- Geniaux, E., Laguerre, G., and Amarger, N. 1993. Comparison of geographically distant populations of *Rhizobium* isolated from root nodules of *Phaseolus vulgaris*. *Mol. Ecol.* 2:295-302.
- Gough, C., and Cullimore, J. 2011. Lipo-chitoooligosaccharide signaling in endosymbiotic plant-microbe interactions. *Mol. Plant-Microbe Interact.* 24:867-878.
- Iriarte, A., Platero, R., Romero, V., Fabiano, E., and Sotelo-Silveira, J. R. 2016. Draft genome sequence of *Cupriavidus* UYMMa02A, a novel beta-Rhizobium species. *Genome Announc.* 4:e01258-16.
- Kim, K. R., and Oh, D. K. 2013. Production of hydroxy fatty acids by microbial fatty acid-hydroxylation enzymes. *Biotechnol. Adv.* 31: 1473-1485.
- Klonowska, A., Chaintreuil, C., Tisseyre, P., Miché, L., Melkonian, R., Ducouso, M., Laguerre, G., Brunel, B., and Moulin, L. 2012. Biodiversity of *Mimosa pudica* rhizobial symbionts (*Cupriavidus taiwanensis*, *Rhizobium mesoamericanum*) in New Caledonia and their adaptation to heavy metal-rich soils. *FEMS Microbiol. Ecol.* 81: 618-635.
- Klonowska, A., Melkonian, R., Miché, L., Tisseyre, P., and Moulin, L. 2018. Transcriptomic profiling of *Burkholderia phymatum* STM815, *Cupriavidus taiwanensis* LMG19424 and *Rhizobium mesoamericanum* STM3625 in response to *Mimosa pudica* root exudates illuminates the molecular basis of their nodulation competitiveness and symbiotic evolutionary history. *BMC Genomics* 19: Article 105.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35:1547-1549.
- Le, S. Q., and Gascuel, O. 2008. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25:1307-1320.
- Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J. C., and Dénarié, J. 1990. Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* 344:781-784.
- Liu, C.-W., and Murray, J. D. 2016. The role of flavonoids in nodulation host-range specificity: An update. *Plants* 5:33.
- Liu, X. Y., Wu, W., Wang, E. T., Zhang, B., Macdermott, J., and Chen, W. X. 2011. Phylogenetic relationships and diversity of β -rhizobia associated with *Mimosa* species grown in Sishuangbanna, China. *Int. J. Syst. Evol. Microbiol.* 61:334-342.
- Lorquin, J., Lortet, G., Ferro, M., Mear, N., Dreyfus, B., Prome, J. C., and Boivin, C. 1997. Nod factors from *Sinorhizobium saheli* and *S. teranga* bv. *sesbaniae* are both arabinosylated and fucosylated, a structural feature specific to *Sesbania rostrata* symbionts. *Mol. Plant-Microbe Interact.* 10:879-890.
- Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczygłowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N., and Stougaard, J. 2003. A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* 425:637-640.
- Marchetti, M., Capela, D., Glew, M., Cruveiller, S., Chane-Woon-Ming, B., Gris, C., Timmers, T., Poinot, V., Gilbert, L. B., Heeb, P., Médigue, C., Batut, J., and Masson-Boivin, C. 2010. Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biol.* 8:e1000280.
- Marchetti, M., Catrice, O., Batut, J., and Masson-Boivin, C. 2011. *Cupriavidus taiwanensis* bacteroids in *Mimosa pudica* Indeterminate nodules are not terminally differentiated. *Appl. Environ. Microbiol.* 77:2161-2164.
- Marchetti, M., Jauneau, A., Capela, D., Remigi, P., Gris, C., Batut, J., and Masson-Boivin, C. 2014. Shaping bacterial symbiosis with legumes by experimental evolution. *Mol. Plant-Microbe Interact.* 27:956-964.
- Martens, M., Dawyndt, P., Coopman, R., Gillis, M., De Vos, P., and Willems, A. 2008. Advantages of multilocus sequence analysis for taxonomic studies: A case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *Int. J. Syst. Evol. Microbiol.* 58:200-214.
- Marx, C. J., and Lidstrom, M. E. 2002. Broad-host-range *cre-lox* system for antibiotic marker recycling in gram-negative bacteria. *Biotechniques* 33: 1062-1067.
- Masson-Boivin, C., and Sachs, J. L. 2018. Symbiotic nitrogen fixation by rhizobia-the roots of a success story. *Curr. Opin. Plant Biol.* 44:7-15.
- Mergaert, P., Van Montagu, M., and Holsters, M. 1997. Molecular mechanisms of Nod factor diversity. *Mol. Microbiol.* 25:811-817.
- Mishra, R. P. N., Tisseyre, P., Melkonian, R., Chaintreuil, C., Miché, L., Klonowska, A., Gonzalez, S., Bena, G., Laguerre, G., and Moulin, L. 2012. Genetic diversity of *Mimosa pudica* rhizobial symbionts in soils of French Guiana: Investigating the origin and diversity of *Burkholderia phymatum* and other beta-rhizobia. *FEMS Microbiol. Ecol.* 79:487-503.
- Moulin, L., Klonowska, A., Caroline, B., Booth, K., Vriezen, J. A. C., Melkonian, R., James, E. K., Young, J. P. W., Bena, G., Hauser, L., Land, M., Kyrpides, N., Bruce, D., Chain, P., Copeland, A., Pitluck, S., Woyke, T., Lizotte-Waniewski, M., Bristow, J., and Riley, M. 2014. Complete genome sequence of *Burkholderia phymatum* STM815(T), a broad host range and efficient nitrogen-fixing symbiont of *Mimosa* species. *Stand. Genomic Sci.* 9:763-774.
- Moulin, L., Mornico, D., Melkonian, R., and Klonowska, A. 2013. Draft genome sequence of *Rhizobium mesoamericanum* STM3625, a nitrogen-fixing symbiont of *Mimosa pudica* isolated in French Guiana (South America). *Genome Announc.* 1:e00066-12.
- Moulin, L., Munive, A., Dreyfus, B., and Boivin-Masson, C. 2001. Nodulation of legumes by members of the β -subclass of Proteobacteria. *Nature* 411:948-950.
- Mulligan, J. T., and Long, S. R. 1985. Induction of *Rhizobium meliloti nodC* expression by plant exudate requires nodD. *Proc. Natl. Acad. Sci. U.S.A.* 82:6609-6613.
- Nandasena, K. G., O'Hara, G. W., Tiwari, R. P., and Howieson, J. G. 2006. Rapid *in situ* evolution of nodulating strains for *Biserrula pelecinus* L. through lateral transfer of a symbiosis island from the original mesorhizobial inoculant. *Appl. Environ. Microbiol.* 72:7365-7367.
- Nelson, M. S., and Sadowsky, M. J. 2015. Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. *Front. Plant Sci.* 6:491.
- Nick, G., de Lajudie, P., Eardly, B. D., Suomalainen, S., Paulin, L., Zhang, X., Gillis, M., and Lindström, K. 1999. *Sinorhizobium arboris* sp. nov. and *Sinorhizobium kostense* sp. nov., isolated from leguminous trees in Sudan and Kenya. *Int. J. Syst. Bacteriol.* 49:1359-1368.
- Oldroyd, G. E. D., and Downie, J. A. 2008. Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annu. Rev. Plant Biol.* 59:519-546.
- Oldroyd, G. E. D., Murray, J. D., Poole, P. S., and Downie, J. A. 2011. The rules of engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* 45:119-144.
- Olthoorn, M. M. A., López-Lara, I. M., Petersen, B. O., Bock, K., Haverkamp, J., Spaink, H. P., and Thomas-Oates, J. E. 1998. Novel

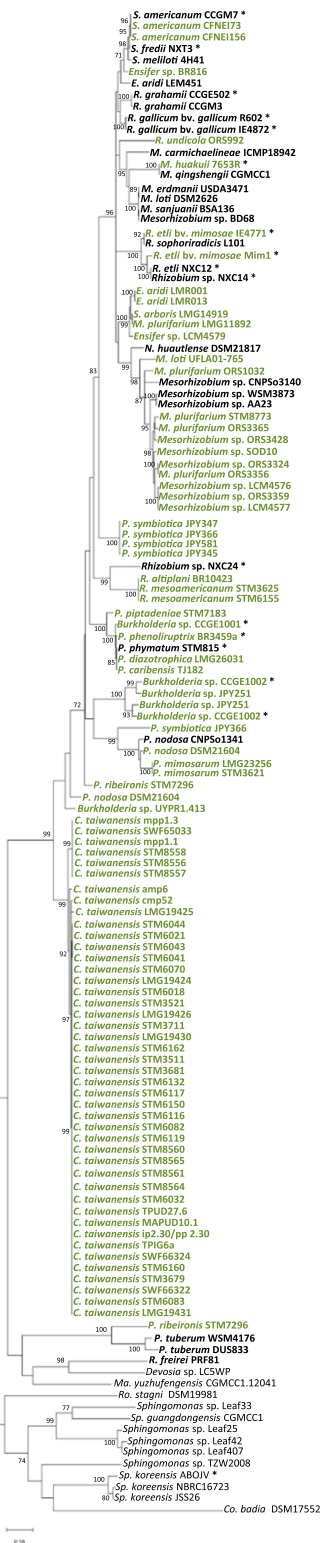
- branched nod factor structure results from α -(1→3) fucosyl transferase activity: The major lipo-chitin oligosaccharides from *Mesorhizobium loti* strain NZP2213 bear an α -(1→3) fucosyl substituent on a nonterminal backbone residue. *Biochemistry* 37:9024-9032.
- Ormeño-Orrillo, E., Menna, P., Almeida, L. G. P., Ollero, F. J., Nicolás, M. F., Pains Rodrigues, E., Shigeyoshi Nakatani, A., Silva Batista, J. S., Oliveira Chueire, L. M., Souza, R. C., Ribeiro Vasconcelos, A. T., Megías, M., Hungria, M., and Martínez-Romero, E. 2012. Genomic basis of broad host range and environmental adaptability of *Rhizobium tropici* CIAT 899 and *Rhizobium* sp. PRF 81 which are used in inoculants for common bean (*Phaseolus vulgaris* L.). *BMC Genomics* 13: Article 735.
- Peck, M. C., Fisher, R. F., and Long, S. R. 2006. Diverse flavonoids stimulate NodD1 binding to *nod* gene promoters in *Sinorhizobium meliloti*. *J. Bacteriol.* 188:5417-5427.
- Perret, X., Staehelin, C., and Broughton, W. J. 2000. Molecular basis of symbiotic promiscuity. *Microbiol. Mol. Biol. Rev.* 64:180-201.
- Persson, T., Battenberg, K., Demina, I. V., Vigil-Stenman, T., Vanden Heuvel, B., Pujic, P., Facciotti, M. T., Wilbanks, E. G., O'Brien, A., Fournier, P., Cruz Hernandez, M. A., Mendoza Herrera, A., Médigue, C., Normand, P., Pawlowski, K., and Berry, A. M. 2015. *Candidatus* Frankia Datiscae Dg1, the actinobacterial microsymbiont of *Datisca glomerata*, expresses the canonical *nod* genes *nodABC* in symbiosis with its host plant. *PLoS One* 10:e0127630.
- Quandt, J., and Hynes, M. F. 1993. Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. *Gene* 127:15-21.
- Rasolomampianina, R., Bailly, X., Fetiariison, R., Rabevohitra, R., Béna, G., Ramarason, L., Raheirmandimby, M., Moulin, L., De Lajudie, P., Dreyfus, B., and Avarre, J. C. 2005. Nitrogen-fixing nodules from rose wood legume trees (*Dalbergia* spp.) endemic to Madagascar host seven different genera belonging to α - and β -proteobacteria. *Mol. Ecol.* 14: 4135-4146.
- Remigi, P., Zhu, J., Young, J. P. W., and Masson-Boivin, C. 2016. Symbiosis within symbiosis: Evolving nitrogen-fixing legume symbionts. *Trends Microbiol.* 24:63-75.
- Ribeiro, R. A., Rogel, M. A., López-López, A., Ormeño-Orrillo, E., Barcellos, F. G., Martínez, J., Thompson, F. L., Martínez-Romero, E., and Hungria, M. 2012. Reclassification of *Rhizobium tropici* type A strains as *Rhizobium leucaenae* sp. nov. *Int. J. Syst. Evol. Microbiol.* 62: 1179-1184.
- Rostas, K., Kondorosi, E., Horvath, B., Simoncsits, A., and Kondorosi, A. 1986. Conservation of extended promoter regions of nodulation genes in *Rhizobium*. *Proc. Natl. Acad. Sci. U.S.A.* 83:1757-1761.
- Schlaman, H. R. M., Horvath, B., Vijgenboom, E., Okker, R. J. H., and Lugtenberg, B. J. J. 1991. Suppression of nodulation gene expression in bacteroids of *Rhizobium leguminosarum* biovar *viciae*. *J. Bacteriol.* 173: 4277-4287.
- Schultze, M., Quiclet-Sire, B., Kondorosi, E., Virelizer, H., Glushka, J. N., Endre, G., Géro, S. D., and Kondorosi, A. 1992. *Rhizobium meliloti* produces a family of sulfated lipooligosaccharides exhibiting different degrees of plant host specificity. *Proc. Natl. Acad. Sci. U.S.A.* 89: 192-196.
- Shanklin, J., Whittle, E., and Fox, B. G. 1994. Eight histidine residues are catalytically essential in a membrane associated iron enzyme, stearyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* 33:12787-12794.
- Snoeck, C., Luyten, E., Poinso, V., Savagnac, A., Vanderleyden, J., and Promé, J. C. 2001. *Rhizobium* sp. BR816 produces a complex mixture of known and novel lipochitoooligosaccharide molecules. *Mol. Plant-Microbe Interact.* 14:678-684.
- Soria-Díaz, M. E., Rodríguez-Carvajal, M. A., Tejero-Mateo, P., Espartero, J. L., Morón, B., Sousa, C., Megías, M., Thomas-Oates, J., and Gil-Serrano, A. M. 2006. Structural determination of the Nod factors produced by *Rhizobium gallicum* bv. *gallicum* R602. *FEMS Microbiol. Lett.* 255:164-173.
- Sprouffske, K., and Wagner, A. 2016. Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinf.* 17: Article 172.
- Sullivan, J. T., Patrick, H. N., Lowther, W. L., Scott, D. B., and Ronson, C. W. 1995. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. *Proc. Natl. Acad. Sci. U.S.A.* 92:8985-8989.
- Sullivan, J. T., and Ronson, C. W. 1998. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proc. Natl. Acad. Sci. U.S.A.* 95:5145-5149.
- Velázquez, E., Igual, J. M., Willems, A., Fernández, M. P., Muñoz, E., Mateos, P. F., Abril, A., Toro, N., Normand, P., Cervantes, E., Gillis, M., and Martínez-Molina, E. 2001. *Mesorhizobium chacoense* sp. nov., a novel species that nodulates *Prosopis alba* in the Chaco Arido region (Argentina). *Int. J. Syst. Evol. Microbiol.* 51:1011-1021.
- Wang, E. T., Rogel, M. A., García-de los Santos, A., Martínez-Romero, J., Cevallos, M. A., and Martínez-Romero, E. 1999. *Rhizobium etli* bv. *mimosae*, a novel biovar isolated from *Mimosa affinis*. *Int. J. Syst. Bacteriol.* 49:1479-1491.
- Wilson, K. 1987. Preparation of genomic DNA from bacteria. Pages 241-245 in: *Current Protocols in Molecular Biology*. F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, eds. John Wiley & Sons, New York, NY, U.S.A.
- Yang, G. P., Debellé, F., Savagnac, A., Ferro, M., Schiltz, O., Maillet, F., Promé, D., Treilhou, M., Vialas, C., Lindstrom, K., Dénarié, J., and Promé, J. C. 1999. Structure of the *Mesorhizobium huakuii* and *Rhizobium galegae* Nod factors: A cluster of phylogenetically related legumes are nodulated by rhizobia producing Nod factors with α , β -unsaturated *N*-acyl substitutions. *Mol. Microbiol.* 34: 227-237.
- Yates, R. J., Howieson, J. G., Reeve, W. G., and O'Hara, G. W. 2011. A re-appraisal of the biology and terminology describing rhizobial strain success in nodule occupancy of legumes in agriculture. *Plant Soil* 348: 255-267.
- Zurdo-Piñero, J. L., Velázquez, E., Lorite, M. J., Brelles-Mariño, G., Schröder, E. C., Bedmar, E. J., Mateos, P. F., and Martínez-Molina, E. 2004. Identification of fast-growing rhizobia nodulating tropical legumes from Puerto Rico as *Rhizobium gallicum* and *Rhizobium tropici*. *Syst. Appl. Microbiol.* 27:469-477.



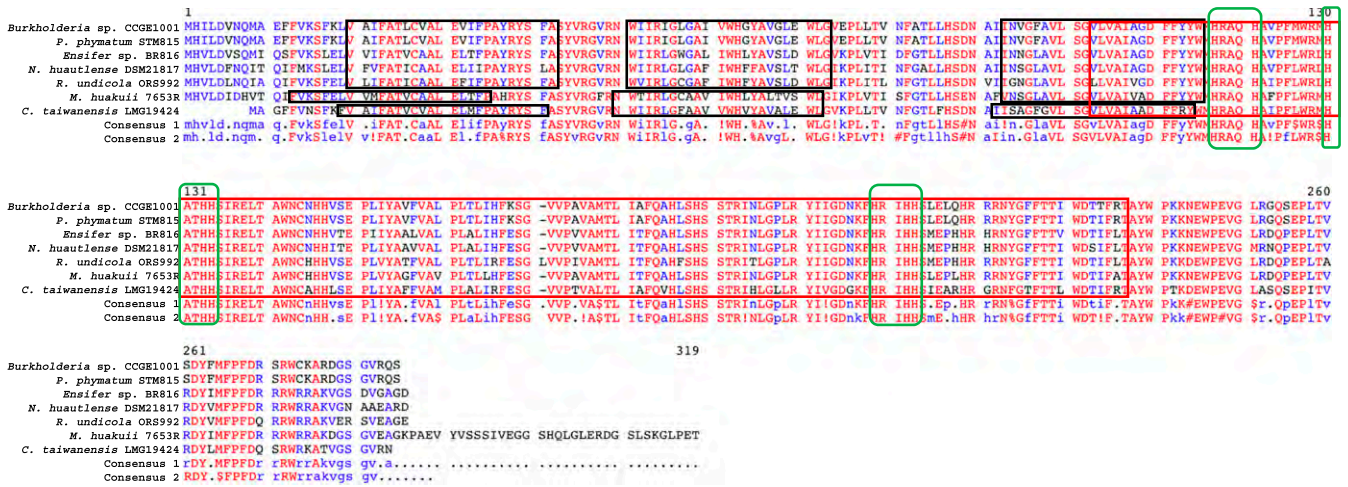
Supplementary Figure S1. Expression of *noeM* in free-living cultures. Wild-type (WT) and mutant ($\Delta nodD$) strains of *C. taiwanensis* bearing a mCherry fusion were cultivated in MPS minimal medium in presence (+L) or absence (-L) of 10 μ M luteolin. mCherry fluorescence and OD (initial OD=0.1) were monitored over 24h. *PpsbA*-mCherry and *PnifH*-mCherry were used as positive and negative controls, respectively. Stationary phase was reached between 10 and 12h after medium inoculation. Treatments with the same letter are not significantly different ($p > 0.05$, ANOVA) considering the logistic AUC.



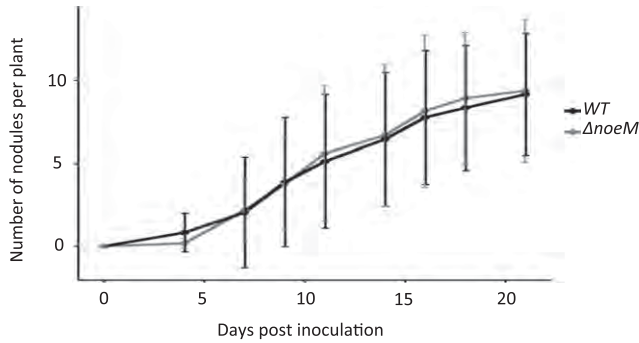
Supplementary Figure S2. A) UPLC-MS/MS spectra of *C. taiwanensis* Nod factors. The MS/MS spectra of the molecular ions $[M+H]^+$: m/z 1393 and m/z 1423.7 are presented on the top and bottom respectively. The two identical fragmentation pattern indicate that the structure of the two molecules is identical from the non-reducing end to the penultimate GINAc subunit. The mass difference affects only the terminal sugar ("reducing end"). **B)** FT-ICR-MS analysis. Left: the isotopic profile of the molecular ion obtained in the negative ionisation mode for the major open oxidized NF. Right: formula deduced from the exact mass observed. **C)** ESI/MS of the methylated sample. After methylation with diazomethane the open-oxidized compound can accept up to two methylations indicating the presence of two carboxylic acids.



Supplementary Figure S3. Phylogenetic tree of genes encoding protein sequences homologous to *C. taiwanensis* NoeM. Maximum likelihood phylogenetic tree based on 131 sequences with 2197 nucleotidic positions inferred by the MEGA X program. Numbers at internal nodes correspond to bootstrap support values (in %), obtained in the analysis of 1000 replicates using the same program and model of sequence evolution. Only bootstraps above 70 are indicated. Truncated sequences have been excluded from the analysis. Strains in bold are rhizobia, strains in green were isolated from a plant belonging to the Mimoseae tribe and * indicates assembled genomes. *C.*, *Cupriavidus*. *Co.*, *Comamonas*. *E.*, *Ensifer*. *Ma.*, *Massilia*. *M.*, *Mesorhizobium*. *P.*, *Parabrucella*. *R.*, *Rhizobium*. *Ro.*, *Roseomonas*. *S.*, *Sinorhizobium*. *Sp.*, *Sphingomonas*.



Supplementary Figure S4. Alignment of NoeM proteins. Multiple sequence alignment with hierarchical clustering performed on protein sequences with Multalin version 5.4.1. *P.*, *Paraburkholderia*; *R.*, *Rhizobium*; *M.*, *Mesorhizobium*; *C.*, *Cupriavidus*. Consensus 1 was obtained from the alignment of the 7 strains above and consensus 2 was obtained from the alignment of all full NoeM sequences in rhizobia. Black and red boxes represent the transmembrane and fatty acid hydroxylase domains, respectively, as predicted using TMHMM Server v.2.0 and PFAM 32.0. HXXXH and HXXXHH motifs are highlighted by green boxes. Amino acid positions are given on the top of the sequences. Consensus symbols: ! is anyone of IV; \$ is anyone of LM; % is anyone of FY; # is anyone of NDQEBZ.



Supplementary Figure S5. Nodulation kinetics. *M. pudica* plants were single-inoculated with either the wild-type (WT) or a $\Delta noeM$ mutant *C. taiwanensis*. *C. taiwanensis* "WT" strain is the Str^R derivative of LMG19424 (CBM832). Data are the results of at least 3 independent experiments. No significant differences were observed between the number of nodules formed by both strains ($p > 0.05$, ANOVA).

Supplementary Table 1. Lipochitooligosaccharide molecules produced with an abundance over 2% by *C. taiwanensis* LMG19424 and *R. solanacearum* GMI1000 derivatives.

strain	MH+	Nomenclature
<i>C. taiwanensis</i> CBM832 and LMG19424	1336.6	LCO V(C18:1, S)
	1367.6	LCO V(C16:0, NMe, Cb, S)
	1379.7	LCO V(C18:1, Cb, S)
	1380.6	LCO open-oxidized (C18:1, NMe, S)
	1393.6	LCO V(C18:1, NMe, Cb, S)
	1395.7	LCO V(C18:0, NMe, Cb, S)
	1397.6	LCO open-oxidized (C16:0, NMe, Cb, S)
	1409.7	LCO open-oxidized (C18:1, Cb, S)
	1423.7	LCO open-oxidized (C18:1, NMe, Cb, S)
<i>C. taiwanensis</i> CBM832ΔnoeM	1336.6	LCO V(C18:1, S)
	1338.7	LCO V(C18:0, S)
	1367.6	LCO V(C16:0, NMe, Cb, S)
	1379.7	LCO V(C18:1, Cb, S)
	1393.6	LCO V(C18:1, NMe, Cb, S)
	1395.7	LCO V(C18:0, NMe, Cb, S)
<i>C. taiwanensis</i> CBM832ΔnoeM::<i>noeM</i>	1336.6	LCO V(C18:1, S)
	1367.6	LCO V(C16:0, NMe, Cb, S)
	1379.7	LCO V(C18:1, Cb, S)
	1380.6	LCO open-oxidized (C18:1, NMe, S)
	1393.6	LCO V(C18:1, NMe, Cb, S)
	1395.7	LCO V(C18:0, NMe, Cb, S)
	1397.6	LCO open-oxidized (C16:0, NMe, Cb, S)
	1409.7	LCO open-oxidized (C18:1, Cb, S)
	1423.7	LCO open-oxidized (C18:1, NMe, Cb, S)
<i>R. solanacearum</i> GMI1000pRalta	1336.6	LCO V(C18:1, S)
	1367.6	LCO V(C16:0, NMe, Cb, S)
	1379.7	LCO V(C18:1, Cb, S)
	1380.6	LCO open-oxidized (C18:1, NMe, S)
	1393.6	LCO V(C18:1, NMe, Cb, S)
	1395.7	LCO V(C18:0, NMe, Cb, S)
	1397.6	LCO open-oxidized (C16:0, NMe, Cb, S)
	1409.7	LCO open-oxidized (C18:1, Cb, S)
	1423.7	LCO open-oxidized (C18:1, NMe, Cb, S)
<i>R. solanacearum</i> GMI1000pRaltaΔnoeM	1336.6	LCO V(C18:1, S)
	1338.7	LCO V(C18:0, S)
	1367.6	LCO V(C16:0, NMe, Cb, S)
	1379.7	LCO V(C18:1, Cb, S)
	1393.6	LCO V(C18:1, NMe, Cb, S)
1395.7	LCO V(C18:0, NMe, Cb, S)	

Nod factor nomenclature is as previously described (Roche et al. 1991, Cell 67:1131-1143). NMe, N-methyl; S, sulfate, Cb, carbamoyl.

Supplementary Table S3. Primers used in this study

Forward primer name	Forward primer sequence 5'-3'	Restriction site	Reverse primer name	Reverse primer sequence 5'-3'	Restriction site	Amplified region	Product length
oCBM3269	GCtctagaTAACCCGTTTGCCGTTATG	XbaI	oCBM3271	CGggatccGGCGAATGATCCAATTTG	BamHI	<i>C. taiwanensis</i> noeM upstream region	494 bp
oCBM3272	CGggatccTCGCAAGCCAGTCTGAACCG	BamHI	oCBM3274	CCGgaattcGCTCAAGTGGCTGGGCCGTT	EcoRI	<i>C. taiwanensis</i> noeM downstream region	519 bp
oCBM3315	GCCCCATGCGCTGACCAAC		oCBM3316	AGAATGCCAACCTTCGCCCGC		Deleted noeM region in CBM2856	1284 bp
oCBM3279	cctaggTGTCGATACTGTCAGGATCG	AvrII	oCBM3430	GactagtCCTTAATTCGTACGCCACTT	SpeI	<i>C. taiwanensis</i> noeM and its promoter region	1247 bp
oCBM3279	cctaggTGTCGATACTGTCAGGATCG	AvrII	oCBM3165	ggtagcTGATTTACATCAAAAAGGTGC	Acc65I	<i>C. taiwanensis</i> noeM promoter region (404 bp)	415 bp
oCBM2739	AAGGAAATCTTCGAGCAGCC		oCBM2740	CGCGTCTGTTTCATGTTGTCT		noeM inserted in the <i>glmS-Ralta_A206</i> intergenic region	3410 bp
oCBM2617	cctaggTAAACAACGGCTAAGTCGAC	AvrII	oCBM2618	ggtagcGTGATACTCCATTTTGAATTG	Acc65I	<i>C. taiwanensis</i> nifH promoter region(for fusion with mCherry)	560 bp
oCBM3365	GCtctagaCGTGGTTTGACTTTGGCAT	XbaI	oCBM3366	CGggatccAATGATCGAGAGCTGGACGT	BamHI	<i>C. taiwanensis</i> nodD upstream region	532 bp
oCBM3367	CGggattcAGTTTGGCTAGTGTGACGC	BamHI	oCBM3368	CCGgaattcATCTGTCCAACCTGTCTCA	EcoRI	<i>C. taiwanensis</i> nodD downstream region	610 bp
oCBM3358	GCTCTAGACGCCAAGACCGTATCGACAA		oCBM3363	CCGGAATTCAGCTAGATATCCACCCGGCT		Deleted nodD region in CBM2927	1630 bp
oCBM3388	CCGgaattcATAACCCGTTTGCCGTTATG	EcoRI	oCBM3370	GGggatccGGCGAATGATCCAATTTG	Acc65I	<i>R. solanacearum</i> noeM upstream region	496 bp
oCBM3371	CccgctcTCGCAAGCCAGTCTGAACCG	SacII	oCBM3372	gagctcGCTCAAGTGGCTGGGCCGTT	SacI	<i>R. solanacearum</i> noeM downstream region	515 bp
oCBM3315	GCCCCATGCGCTGACCAAC		oCBM3316	AGAATGCCAACCTTCGCCCGC		<i>R. solanacearum</i> noeM interrupted by insertion of kanamycin	2613 bp
oCBM3812	GCGTGGTGATCGATTCTCC		oCBM3813	CGTGAAGGGGACATGGAATTC		<i>hisB</i> fragment	56 bp
oCBM3814	CGAGATCATCAACCCGAACC		oCBM3815	GCTCGACCTTGATCTGCATG		<i>rpoA</i> fragment	77 bp
oCBM1513	TCGAAGGAATGAATGCAGGT		oCBM1514	GCCAAGACCGTATCGACAAT		<i>nodB</i> fragment	100 bp
oCBM1515	TGGCTCTGAATGGTTAGGG		oCBM1516	GCGACTAGAACGCTGAAAG		<i>noeM</i> fragment	118 bp
St_desF1	TTTCCNGCTTAYCGATACAG		St_desR2	AGICCIACYTCNGGCCAYTC		<i>noeM</i> internal fragment in <i>rhizobia</i>	650 bp

Chapitre IV : Résultats complémentaires

Dans ce chapitre, je vais présenter, d'une part les analyses complémentaires qui ont été réalisées pour mieux comprendre le rôle de *noeM* chez les rhizobia et les *Ralstonia* issus de l'expérience d'évolution, et d'autre part les analyses préliminaires de l'infection racinaire de *M. pudica* par *C. taiwanensis* et des clones de *Ralstonia* présentant des niveaux d'infection différents.

IV.1. Analyses complémentaires du gène *noeM*

Grâce aux travaux réalisés sur *C. taiwanensis* et autres rhizobia nous avons pu déterminer que NoeM appartient à un groupe de protéines spécifiques aux rhizobia. Nous avons montré que *noeM* est sous contrôle de la régulation *nod* et qu'il est impliqué dans la synthèse d'un FN atypique dont le sucre réducteur est ouvert et oxydé. Néanmoins, plusieurs questions restent en suspens, notamment le rôle biochimique de cette enzyme mais aussi sa fonctionnalité spécifique par rapport aux homologues avec une identité inférieure à 44%.

Par ailleurs, nous n'avons pas pu démontrer le rôle de *noeM* dans l'infection chez *C. taiwanensis* contrairement aux résultats obtenus avec les clones de la lignée R de l'expérience d'évolution. Ces résultats indiquent un impact de *noeM* dépendant du background génomique des souches. Nous avons donc voulu évaluer dans quel background génomique de *R. solanacearum* cet impact se manifestait.

IV.1.1 Spécificité fonctionnelle de *noeM*

L'étude phylogénétique de NoeM de *C. taiwanensis* nous a permis d'identifier deux groupes de protéines : le groupe I, avec une identité protéique inférieure à 44% à pRalta_0471 et le groupe II qui ne contient que des rhizobia est présentant une identité supérieure à 70% avec NoeM. Afin d'évaluer si seules les protéines du groupe II sont capables d'agir sur les NFs, nous avons décidé de cloner un homologue du groupe I et du groupe II chez le mutant $\Delta noeM$ de *C. taiwanensis*. Pour cela, nous avons sélectionné la souche de *Paraburkholderia tuberum* WSM4176, rare rhizobium dont le gène homologue de *noeM* encode une protéine du groupe I. Une étude des NFs synthétisés par *Rhizobium gallicum* sv. *gallicum* R602 a démontré que cette souche produit des N-glucosaminitol (Soria-Díaz et al. 2006), possédant un homologue de *noeM* du groupe II, cette souche était pour nous une candidate idéale pour faire office de contrôle positif. Ces gènes ont été synthétisés par la compagnie Twist Bioscience, vérifié par séquençage et clonés dans pBBR1-MCS3 (Kovach et al. 1995). J'ai remplacé le gène *noeM* de *C. taiwanensis* par chacune de ces séquences, générant deux souches de CBM832 possédant un homologue de *noeM* en aval de la *nod*-box (**Table IV.1.1.1**) et les constructions ont été vérifiées par séquençage (**Figure IV.1.1.1**). Des cultures bactériennes ont été produites

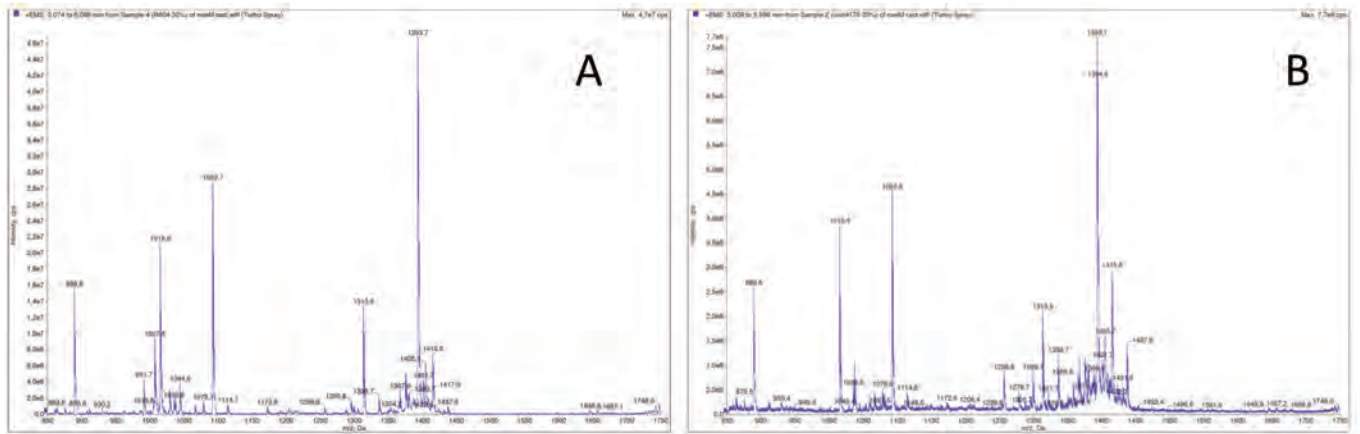


Figure IV.1.1.2: Spectre UPLC-MS des facteur Nod produit par les mutants de *C. taiwanensis*. A) *C. taiwanensis* CBM832 Δ *noeM* pBBR1::*pnoeM*::*noeMR602* et B). *C. taiwanensis* CBM832 Δ *noeM* pBBR1::*pnoeM*::*noeMWSM4176*. La fraction SPE en phase inverse éluée avec 50% d'acétonitrile a été soumise à une analyse par spectrométrie de masse en tandem sur un système Q-Trap4500 couplé à un dispositif UPLC. Le spectre présenté est la somme des spectres obtenus sur le domaine de temps de rétention NF.

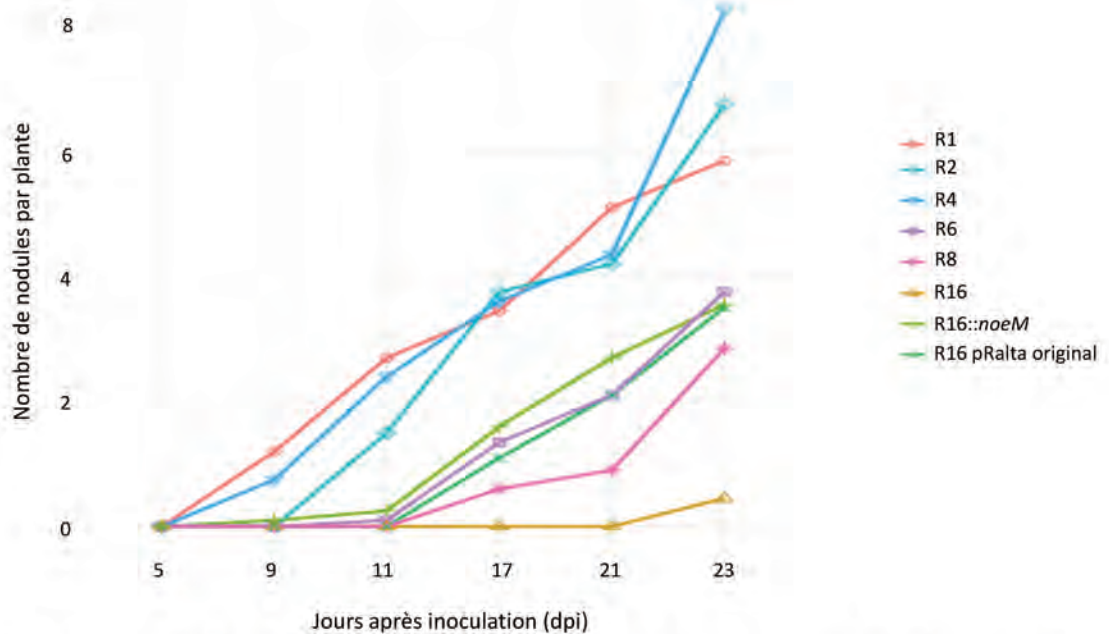


Figure IV.1.2.1: Cinétique de nodulation des clones de la lignée R. 20 plantules de *M. pudica* ont été inoculées avec différents les clones R1, R2, R4, R6, R8, R16 de la lignée R ainsi que les souches R16::*noeM* et R16 pRalta original. Les premiers nodules sont apparus entre 5 et 21 dpi en fonction du clone.

comme décrit précédemment (Daubech et al. 2019) et induites à la lutéoline afin d'analyser la structure des FNs produits.

Les FNs synthétisés par la souche CBM832 Δ noeM ::*noe*MWSM4176 sont similaires à ceux de CBM832 Δ noeM et ne présentent pas de structure atypique. Cependant, et de façon surprenante, la souche CBM832 Δ noeM ::*noe*M R602 a également produit des FNs identiques à ceux du mutant *noe*M de CBM832 (**Figure IV.1.1.2**). Cette absence de complémentarité ne nous a pas permis de conclure quant à la spécificité de NoeM présente chez les rhizobia du groupe II.

NoeM est une protéine avec un domaine catalytique prédit comme une fatty acid hydroxylase. Or nous avons trouvé que NoeM agit sur la partie sucre de lipochitoooligosaccharides, activité non décrite jusqu'à présent chez cette famille de protéines. Afin d'identifier le rôle précis de cette enzyme dans la structure des FNs nous avons collaboré avec Sébastien Fort directeur de recherche en chimie et biotechnologie des oligosaccharides au Centre de Recherches sur les Macromolécules Végétales (CERMAV). Le premier objectif de l'équipe de S. Fort était de tester *in vitro* l'activité biochimique de l'enzyme. NoeM étant une protéine transmembranaire, comme NodC, son équipe a cloné *noe*M afin de le produire chez *E. coli* C43. Après cultures, les cellules d' *E. coli* C43-NoeM ont été lysées, les culots de fractions membranaires lysées récupérées afin de tester l'activité enzymatique de NoeM *in vitro* par incubation de CO-V (s). L'analyse MALDI-TOF en fin de réaction n'a pas permis de mettre en évidence l'activité de NoeM *in vitro*.

Une seconde méthode, *in vivo*, a ensuite été envisagée. L'objectif était de construire une souche de *E. coli* productrice d'oligosaccharides de chitine où le gène *noe*M a été ajouté et d'observer si les COs normalement produits (CO-V(N^V,S) et CO-V(S)) sont modifiés. Deux souches de *E. coli* DEK ont ainsi été transformées, *E. coli* DEK-*nodCBH* et *E. coli* DEK-*nodCBH-noe*M. Les COs formés par les deux souches ont ensuite été analysés par MALDI-TOF. Les profils de masse ont montré que les sucres produits par les deux souches n'étaient pas modifiés, ne permettant pas de mettre en évidence la présence de sucres réduits ou oxydés.

IV.1.2. Impact de *noe*M chez *R. solanacearum*

La compréhension des mécanismes qui permettent l'infection des cellules du nodule est l'un des objectifs majeurs de l'équipe. Le gène *noe*M a été identifié au cours de l'expérience d'évolution où il joue un rôle crucial dans l'infection (**Figure III.1.1** et **Figure III.1.3**) et la nodulation dans la lignée R (**Figure IV.1.2.1**). Néanmoins, nos travaux sur le modèle de référence *C. taiwanensis*/*M. pudica* ont montré que ce gène n'influence aucunement la capacité du rhizobium à infecter les cellules du nodule. Nous avons donc émis l'hypothèse que l'effet de *noe*M soit plus fort dans un contexte génomique mal adapté à la symbiose. Ainsi, la perte de *noe*M aurait un impact plus important chez un clone ancestral que chez un clone évolué plus adapté. Afin de vérifier cette hypothèse nous avons décidé de faire l'étude fonctionnelle de ce gène chez des souches de *R. solanacearum* portant le pRalta et des

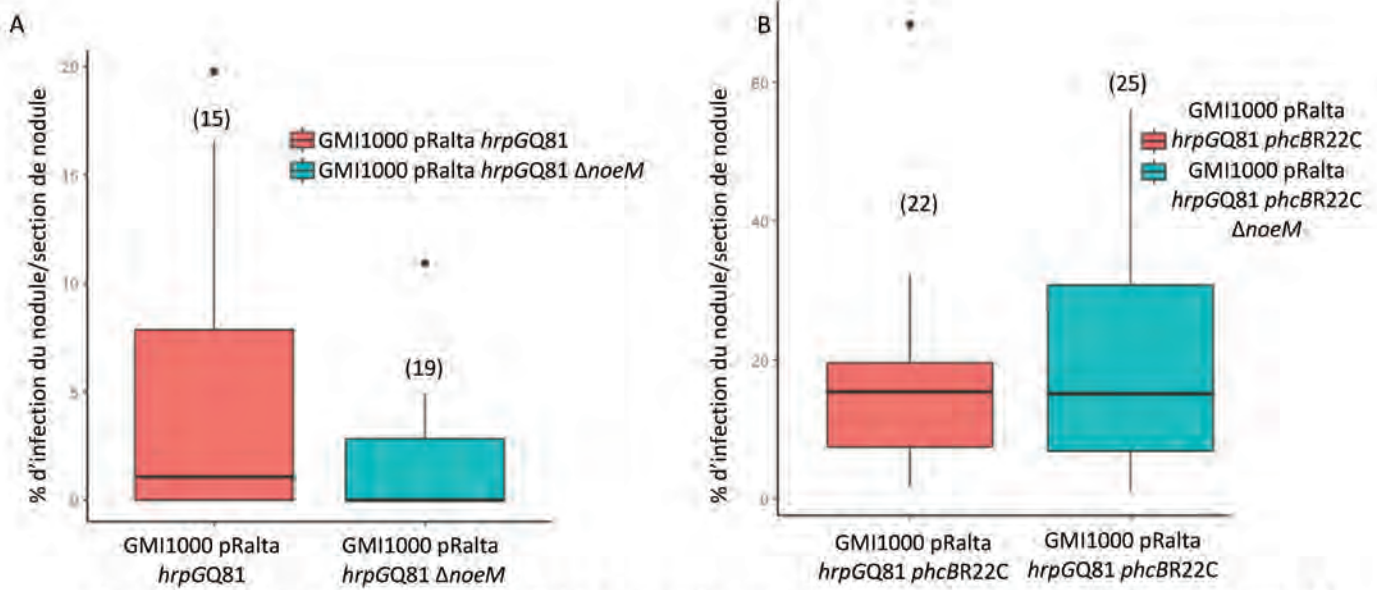


Figure IV.1.2.2: Capacité d'infection intracellulaire du nodule des clones de *R. solanacearum* GMI1000 pRalta *hrpG* et *R. solanacearum* GMI1000 pRalta *hrpG phcB* et de leur mutant *noeM*. Box plot représentant l'aire d'infection par coupe de nodule de GMI1000 pRalta *hrpG* et son mutant *noeM* (A), et de GMI1000 pRalta *hrpG phcB* et son mutant *noeM* (B). Cette analyse a été réalisée une seule fois, le nombre de coupes de nodule utilisées est indiqué entre parenthèses au dessus des box plot.

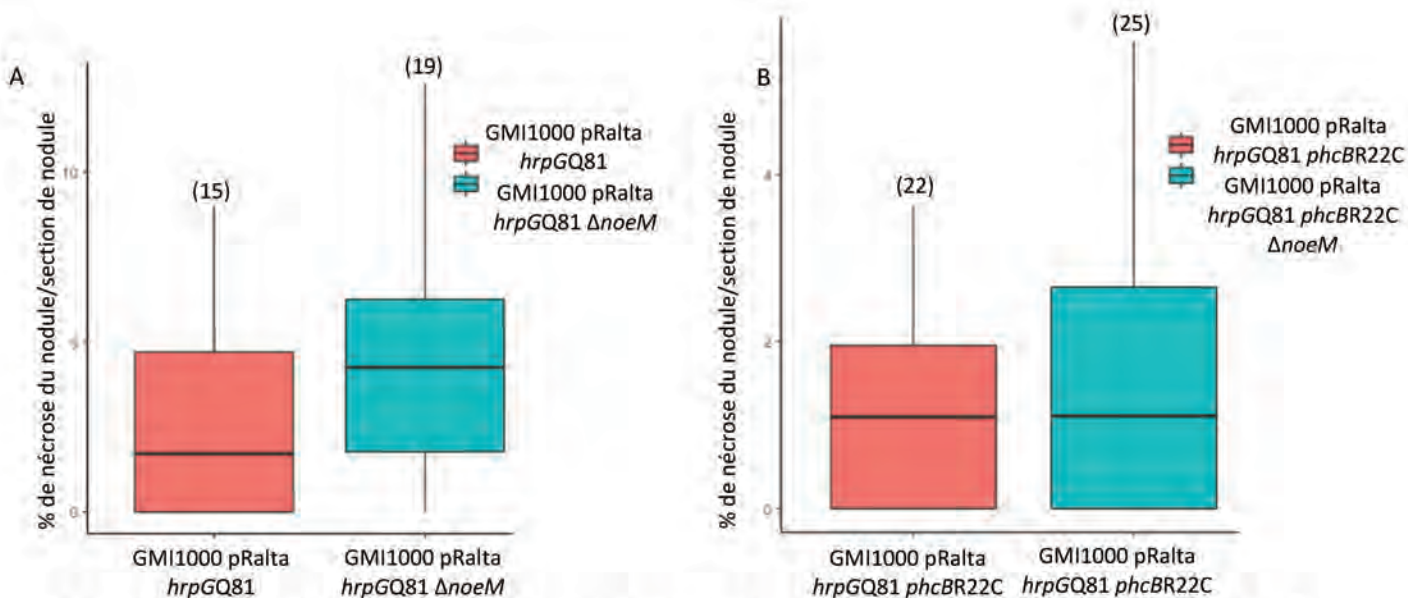


Figure IV.1.2.3: Quantification des nécroses provoquées par les clones de *R. solanacearum* GMI1000 pRalta *hrpG* et *R. solanacearum* GMI1000 pRalta *hrpG phcB* et de leur mutant *noeM*. Box plot représentant l'aire d'infection par coupe de nodule de GMI1000 pRalta *hrpG* et son mutant *noeM* (A), et de GMI1000 pRalta *hrpG phcB* et son mutant *noeM* (B). Cette analyse a été réalisée une seule fois, le nombre de coupes de nodule utilisées est indiqué entre parenthèses au dessus des box plot.

mutations impliquées dans l'amélioration de l'infection qui sont GMI1000pRalta *hrpGQ81* et GMI1000pRalta *hrpGQ81 phcBR22C*. La mutation stop de *hrpGQ81* confère la capacité à noduler et infecter de façon rudimentaire tandis que la mutation *phcBR22C*, placé dans un background *hrpGQ81*, améliore fortement la capacité d'infection intracellulaire. Le gène *noeM* a été délété chez ces deux souches, générant les souches GMI1000pRalta *hrpGQ81ΔnoeM* et GMI1000pRalta *hrpGQ81phcBR22CΔnoeM*. Pour ce faire, les régions 5' et 3' du gène ont été amplifiées par PCR à partir de l'ADN génomique de *C. taiwanensis* en utilisant les primers EcoRI_*noeM* (CCGgaattcATAACCCGTTTGCCGTTATG) et Acc65I_*noeM* (GGggtaccGGCGAATGATCCAATTCG) pour la région en amont du gène et les primers SacII (CccgcggTCGCAAGCCAGTCTGAACCG) et SacI (gagctcGCTCAAGTGGCTGGGCCGTT) pour la région en aval du gène. Ces produits PCR ont ensuite été digérés avec EcoRI/Acc65I et SacI/SacII puis clonés l'un après l'autre dans le plasmide pCM184 (Marx and Lidstrom 2002). Le produit de ligation a été transformé dans la souche de *E. coli* DH5α. Le plasmide a été linéarisé par une digestion avec EcoRI et introduit dans les souches GMI1000 pRalta *hrpGQ81* et GMI1000 pRalta *hrpGQ81phcBR22C* par transformation naturelle comme décrit dans les travaux de Cruz (Cruz et al. 2014). Les transformants ont été sélectionnés par résistance à la kanamycine.

Les phénotypes symbiotiques des 4 mutants ont été évalués sur *M. pudica*. Les plantes ont été cultivées comme décrit précédemment (Marchetti et al. 2010) et inoculées avec 10⁶ bactéries par tube. Afin de mesurer la capacité d'infection intracellulaire des nodules formés par les différents mutants, les nodules formés ont été récoltés à 21 dpi et coupés en sections de 55μm avec un appareil vibratome Leica VT1000S. Les coupes ont ensuite été observées en utilisant un microscope inversé (DMIRBE; Leica) et les images acquises avec un système de caméra CDD (Color Coolview, Photonic Science, Milham, UK). Une analyse quantitative de l'infection et de la nécrose a été réalisée sur les images comme précédemment décrit (Marchetti et al. 2014). Les surfaces d'infection (**Figure IV.1.2.2**) et de nécrose (**Figure IV.1.2.3**) des deux mutants *noeM* se sont avérées être similaires à celles des deux contrôles GMI1000 pRalta *hrpG* et GMI1000 pRalta *hrpG phcB*. Ces résultats indiquent que dans ces backgrounds génomiques la mutation n'affecte pas la capacité infectieuse intracellulaire du nodule, comme nous l'avons constaté avec *C. taiwanensis*. Néanmoins ces résultats ont été obtenus à partir d'un seul réplica biologique.

En parallèle, nous avons comparé les capacités de nodulation, soit la cinétique de nodulation et la compétition pour la nodulation des clones de *R. solanacearum*. Pour comparer les cinétiques de nodulation, des plants de *M. pudica* ont été inoculés avec 10⁶ bactéries par tube pour chacune des souches de GMI1000 pRalta *hrpG phcB* et de son mutant *noeM*. Le mutant *noeM* n'a montré qu'un léger retard de nodulation comparé à la souche sauvage, les premiers nodules du mutant *noeM* apparaissant à partir de 7 dpi alors que la souche GMI1000 pRalta *hrpG phcB* nodule dès 5 dpi (**Figure IV.1.2.4**). Pour analyser la compétition pour la nodulation, les plantes ont été co-inoculées avec GMI1000 pRalta *hrpG* ou GMI1000 pRalta *hrpG phcB* et leur mutant *noeM* respectif (ratio 1/1, 10⁶ bactéries totales par tubes). A 21

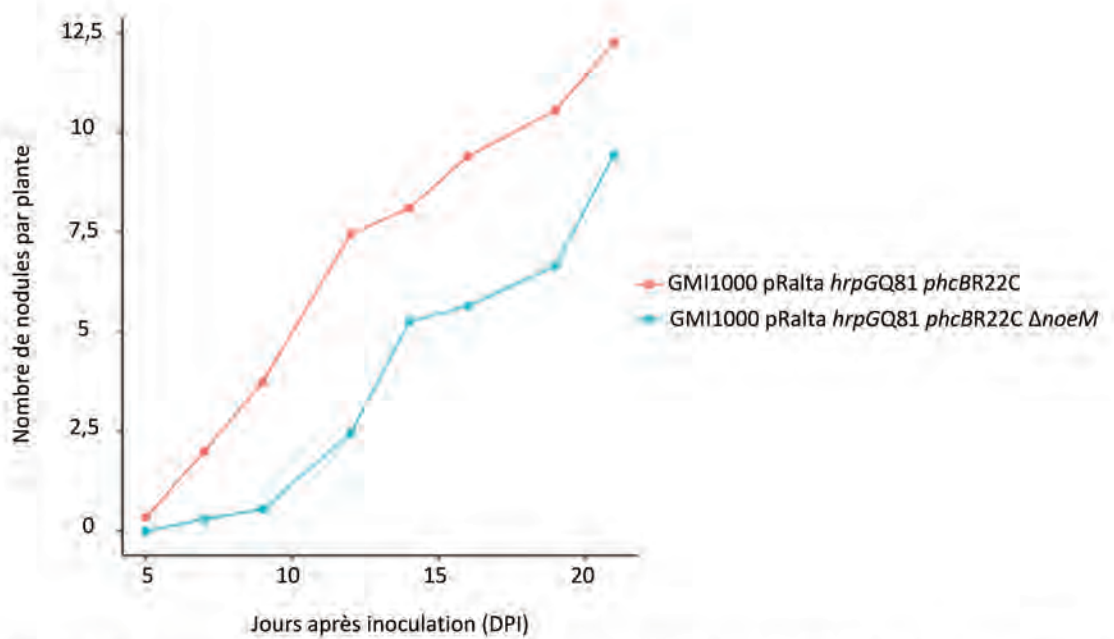


Figure IV.1.2.4: Cinétique de nodulation de GMI1000 pRalta *hrpGQ81 phcBR22C* et de son mutant $\Delta noeM$. Les plantules de *M. pudica* ont été inoculées avec la souche GMI1000 pRalta *hrpGQ81 phcBR22C* ou la souche GMI1000 pRalta *hrpGQ81 phcBR22CΔnoeM*. Les premiers nodules sont apparus à 5 dpi.

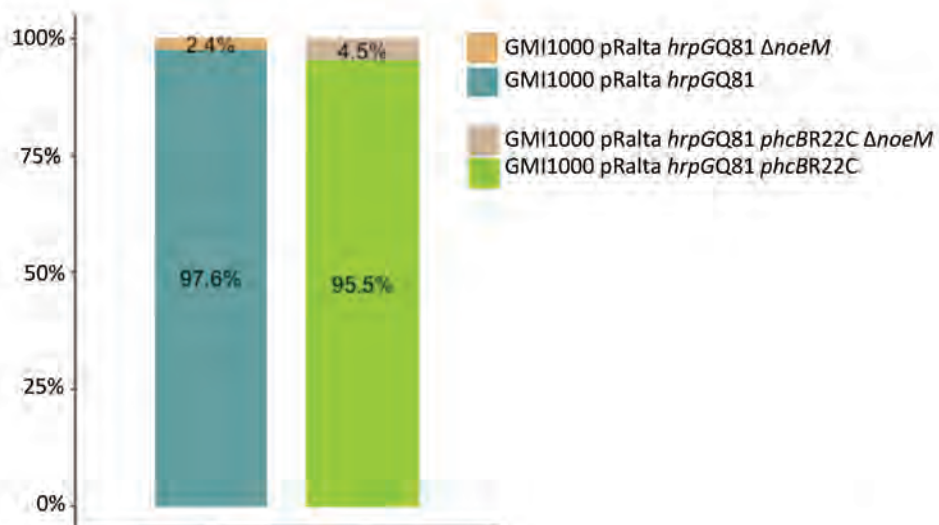


Figure IV.1.2.5: Compétition pour la nodulation de mutants de *R. solanacearum* GMI1000 pRalta. Diagramme à barres empilées des pourcentages de nodules de plants de *M. pudica* co-inoculées avec une mixture (ratio 1/1) de GMI1000 pRalta *hrpGQ81* et GMI1000 pRalta *hrpGQ81 ΔnoeM* et des plantules de *M. pudica* co-inoculées avec une mixture (ratio 1/1) de GMI1000 pRalta *hrpGQ81 phcBR22C* et GMI1000 pRalta *hrpGQ81 phcBR22CΔnoeM*.

dpi, 96 nodules ont été récoltés et stérilisés comme décrit précédemment (Daubech et al. 2017) avant d'être broyés individuellement et déposés sur boîte avec et sans kanamycine, les mutants *noeM* étant les seuls à être résistants à la kanamycine. La majorité des nodules récoltés était infectés par les souches non délétées (entre 95% et 98% de la totalité des nodules) de *noeM* indiquant que le mutant *noeM* est très affecté pour la compétition pour la nodulation comparé à la souche portant l'allèle sauvage (**Figure IV.1.2.5**). Cet effet est plus fort que chez *C. taiwanensis* où les mutants *noeM* représentaient environ 30% des bactéries isolées des nodules.

Nos travaux montrent donc que l'impact la délétion de *noeM* dans la symbiose, et notamment sur l'infection intracellulaire, avec *M. pudica* est dépendant du background génétique. En effet, la délétion de *noeM* affecte clairement la capacité à infecter les nodules du clone évolué R16 (la réintroduction du plasmide symbiotique ou de *noeM* chez R16 restaure un phénotype infectieux (Marchetti et al. 2016)), mais pas de *C. taiwanensis* et des deux souches GMI1000 pRalta *hrpG* et GMI1000 pRalta *hrpG phcB*. Il est probable que l'effet de *noeM* soit en conjonction avec une ou plusieurs autres mutations spécifique à la lignée R.

IV.2.1. Analyse préliminaire des étapes précoces de l'interaction avec *M. pudica*

L'expérience d'évolution a eu comme effet une amélioration spectaculaire de la capacité d'infection des ancêtres nodulants qui étaient peu ou pas infectieux. Les travaux de l'équipe ont permis d'identifier plusieurs mutations responsables de l'amélioration de l'infection cellulaire (**Figure IV.2.1.1**), c'est-à-dire l'augmentation du nombre de cellules du nodule infectées et la diminution du nombre de nécroses (Marchetti et al. 2010; Guan et al. 2013; Marchetti et al. 2014, 2016; Capela et al. 2017). Cette amélioration peut se traduire par une meilleure capacité à envahir les cellules et à persister, ce qui pourrait être le résultat d'une diminution des réactions de défense ou/et une amélioration du phénomène d'endocytose lors du passage entre la sortie du poil absorbant et l'infection des cellules du nodule. Il est possible que l'amélioration de l'infection provienne d'une meilleure capacité à réaliser les étapes précoces et induire le processus développemental conduisant à l'infection du nodule.

Des travaux sur *Medicago sativa* ont montré que le timing développemental du nodule est important et qu'il a une influence sur l'infection (Xiao et al. 2014). Par exemple, lorsque le gène le gène *ipd3* est inactivé chez *M. sativa*, la progression des cordons d'infection est retardée. Ce retard fait que les cordons d'infection ne peuvent pas passer la zone méristématique après divisions périclines des cellules empêchant l'infection des cellules du nodule (Xiao et al. 2014).

Dans l'expérience d'évolution, il s'avère que les différentes mutations identifiées pour améliorer l'infection intracellulaire entraînent également une amélioration de la compétitivité pour la nodulation (**Figure IV.2.1.2**).

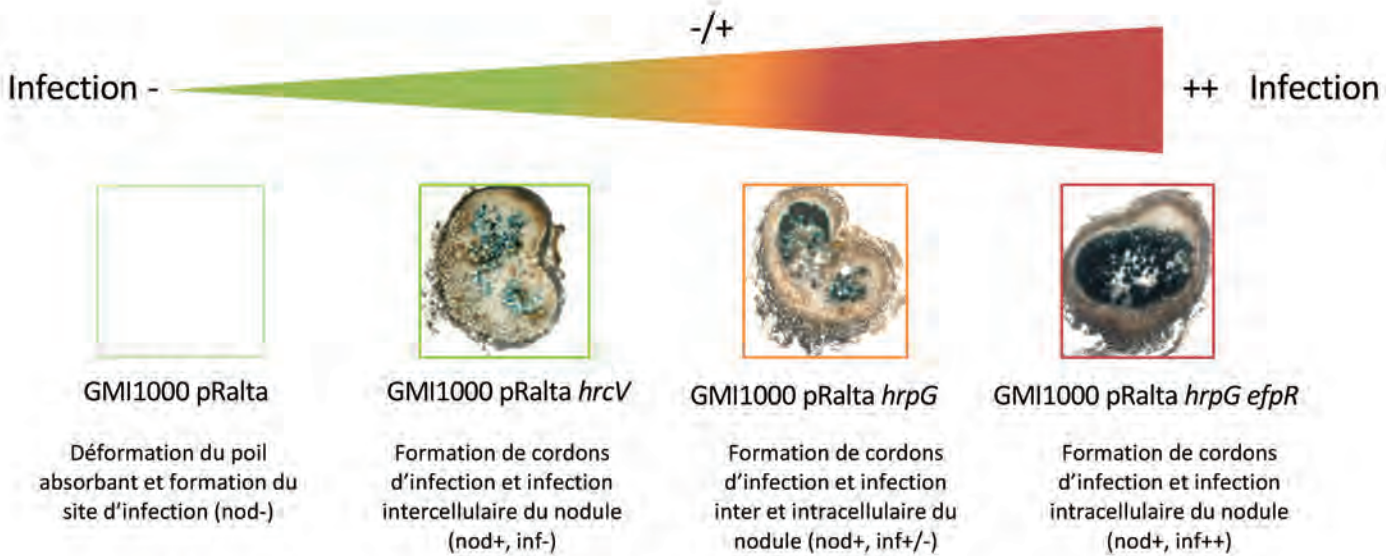


Figure IV.2.1.1: Coupes de nodules et phénotypes infectieux des souches de *R. solanacearum*. Du moins infectieux au plus infectieux *R. solanacearum* GMI1000 pRalta, *R. solanacearum* GMI1000 pRalta *hrcV*, *R. solanacearum* GMI1000 pRalta *hrpG* et *R. solanacearum* GMI1000 pRalta *hrpG/efpR*.

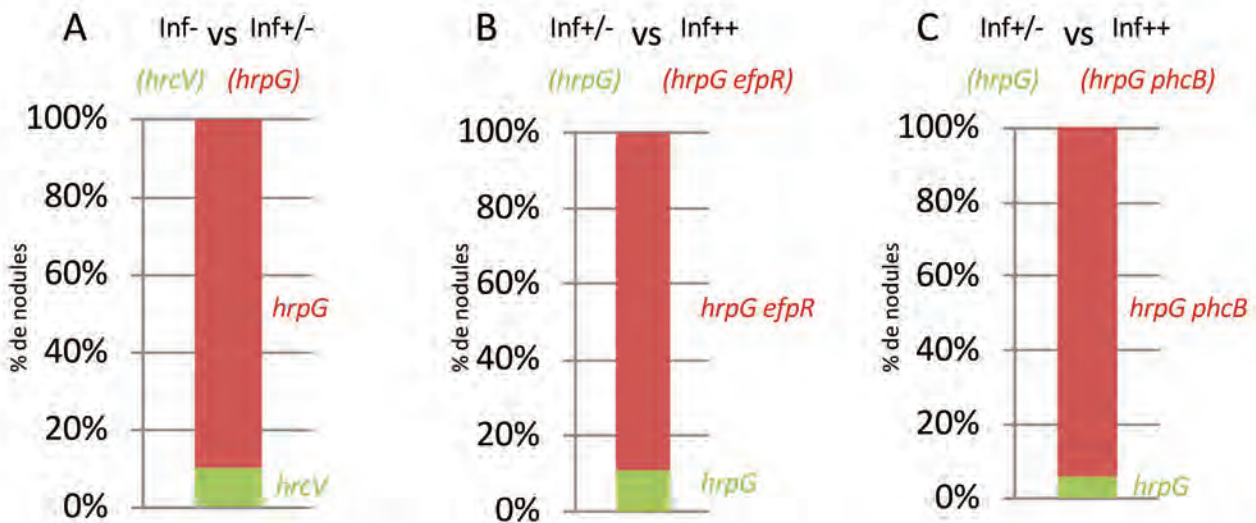


Figure IV.2.1.2: Compétition pour la nodulation. Diagramme à barres empilées des pourcentages de nodules infectés par *R. solanacearum* GMI1000 pRalta *hrcV* ou *R. solanacearum* GMI1000 pRalta *hrpG* (A), *R. solanacearum* GMI1000 pRalta *hrpG*, *R. solanacearum* GMI1000 pRalta *hrpG/efpR* (B) et *R. solanacearum* GMI1000 pRalta *hrpG*, *R. solanacearum* GMI1000 pRalta *hrpG/phcB* (C).

Afin d'évaluer l'impact des mutations sur le processus infectieux, à savoir l'invasion et la persistance des bactéries per se ou bien un effet plus précoce, nous avons initié l'analyse des étapes précoces de l'infection racinaire de *M. pudica* par *C. taiwanensis* et les clones évolués de *R. solanacearum*. Les étapes précoces de l'infection comprennent le moment où les bactéries vont s'accrocher à la paroi végétale du poil absorbant jusqu'au relâchement des bactéries dans les cellules du nodules. Une première étude cytologique a montré que *C. taiwanensis* pénètre les tissus racinaires via les poils absorbants à partir desquels se forment les cordons d'infection qui vont conduire à la libération des bactéries dans le cortex racinaire (Marchetti et al. 2011). Cependant, aucun travaux n'a déterminé comment se déroule l'initiation et la progression du cordon d'infection. Dans cette dernière partie, je vais présenter les résultats préliminaires obtenus avec le rhizobium de référence *C. taiwanensis* ainsi que les clones de *Ralstonia* GMI1000 pRalta, incapable de noduler et son mutant isogénique *hrcV*, capable de former des nodules avec une infection exclusivement intercellulaire. Les plants de *M. pudica* ont été préparés comme décrit précédemment (Marchetti et al. 2010) et cultivés dans des boîtes de Pétri carrées, sur milieu Fahraeus. Pour chaque boîte, les racines ont été inoculées avec un millilitre d'inoculum de chaque souche (10^6 bactéries/ml) portant le gène rapporteur *lacZ* sous contrôle du promoteur du gène *nodB* (Marchetti et al. 2010). Pour l'analyse de la cinétique des étapes précoces du processus infectieux, entre 5 et 9 racines ont été récoltées à 3, 5, 7, 9 et 11 dpi puis fixées au glutaraldehyde 2.5% avec 1% de triton 100X dans du tampon phosphate pendant 30 minutes sous vide. Après rinçage, les racines ont été incubées pendant la nuit avec la solution de coloration 28°C (0.1 M K phosphate [pH 7.4], 2 mM K ferricyanide, 2 mM K ferrocyanide, et 0.08% of X-gal dans du diméthylformamide). Les racines ont été rincées puis analysées au microscope (Zeiss Axiophot I).

Pour chaque racine et pour chaque point de récolte, le nombre de sites d'infection, de cordons d'infection et de nodules ont été dénombrés (**Figure IV.2.1.3**). Dès 3 dpi, on observe l'apparition de sites d'infection (5,71 sites par plante en moyenne) pour *C. taiwanensis* qui augmentent au cours de la cinétique pour atteindre plus de 20 sites d'infection à 7dpi. Au début de la cinétique, le nombre de cordons d'infection croit rapidement passant de 9,6 cordons en moyenne à 3 dpi à un nombre moyen de 21,4 cordons à 7 dpi. A partir de 7dpi, le nombre de cordons d'infection atteint un plateau et varie très peu. Les premiers nodules formés par *C. taiwanensis* apparaissent dès 5 dpi. Bien que variable en fonction des plantes, le nombre de nodules reste globalement le même tout au long de la cinétique. Une activité fixatrice d'azote a été mesuré dans des nodule de *C. taiwanensis* dès 7 dpi (Daubech et al. 2017) ce qui est en accord avec la mise en place du système d'autorégulation de l'infection bien connu dans les symbiose fixatrice d'azote .

Le même type d'analyse a été réalisé de façon préliminaire pour les deux clones de *R. solanacearum* (**Figure IV.2.1.3**). Concernant GMI1000 pRalta, les sites d'infections apparaissent à partir de 5 dpi, soit plus tard qu'avec *C. taiwanensis*. De la même façon, les premiers cordons d'infection ont pu être visualisés à partir de 5 jours mais en très faible nombre avec 3,625 cordons par plante soit plus de 3,5 fois moins que *C. taiwanensis* à la même

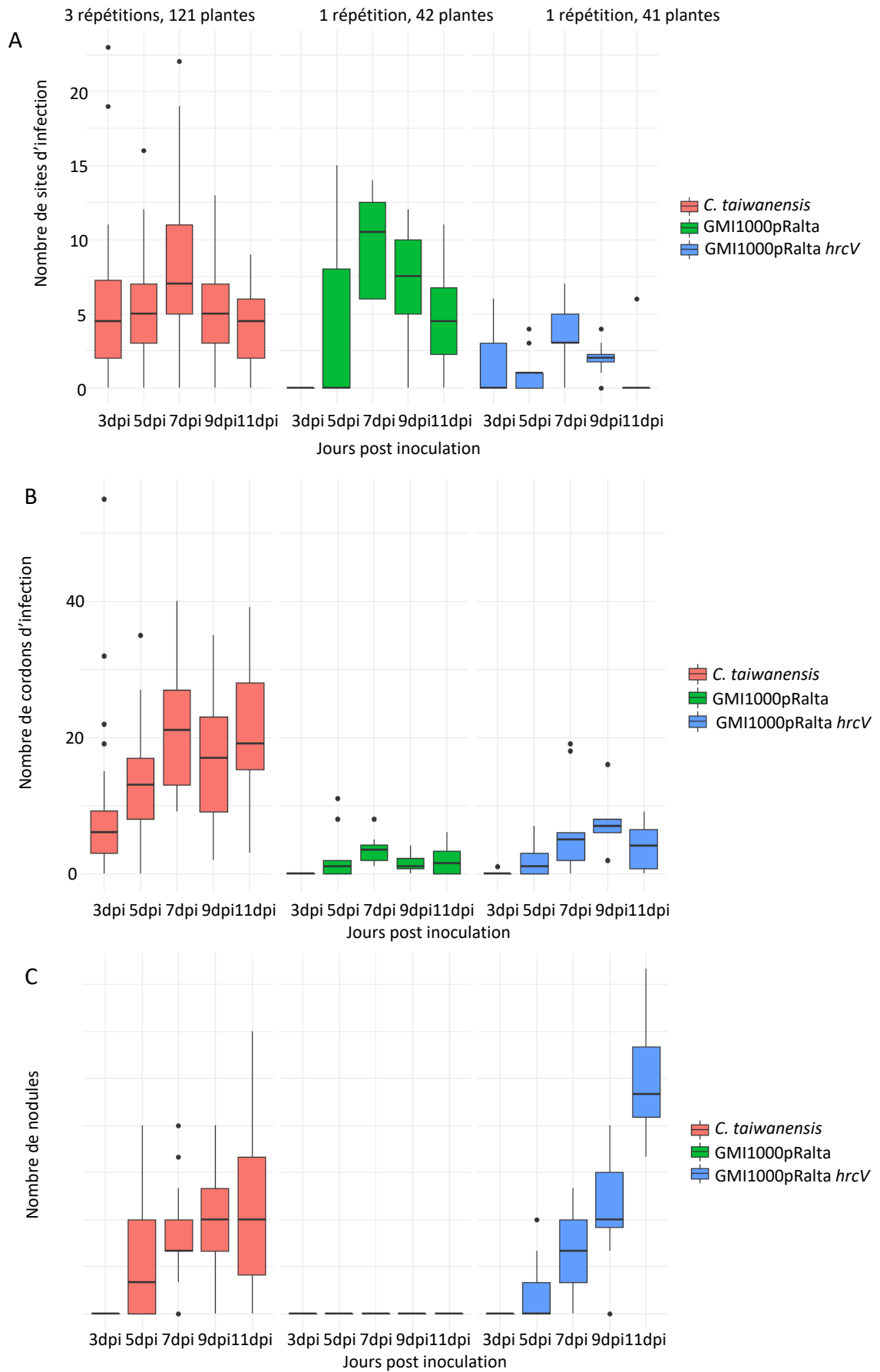


Figure IV.2.1.3: Cinétique de l'infection racinaire de *C. taiwanensis*, *R. solanacearum* GMI1000 pRalta et *R. solanacearum* GMI1000 pRalta *hrcV*. Box plot représentant le nombre de cordons d'infection (A), le nombre de cordons d'infection (B) et le nombre de nodules (C) pour chacune des trois souche .

période. Sans surprise, la souche GMI1000 pRalta n'induit aucun nodule. De façon similaire à *C. taiwanensis*, nous avons pu observer que la souche GMI1000 pRalta *hrcV* est capable d'induire des sites d'infection dès 3 dpi, bien qu'en nombre inférieur avec en moyenne 1,3 sites par racine. Cependant, les cordons d'infection ainsi que les nodules présentent une cinétique décalée par rapport à celle de *C. taiwanensis* car ils apparaissent seulement à 5 dpi. De plus, le nombre de cordons d'infection moyen ne dépassent jamais 7,5 par plante (valeur maximale atteinte à 9 dpi contre 17,2 cordons par plante avec la souche *C. taiwanensis*). Contrairement à *C. taiwanensis*, pour le mutant *hrcV* le nombre de nodules ne cesse d'augmenter au cours de la cinétique, ce qui est en accord avec l'absence de fixation d'azote.

Chapitre V : Conclusions et Discussion générale

L'évolution expérimentale, i.e. la propagation d'organismes sur des centaines voire des milliers de générations dans un environnement contrôlé –souvent le laboratoire-, permet de mieux comprendre comment un organisme s'adapte à des conditions environnementales spécifiques. C'est aussi un outil qui permet de mieux comprendre et d'explorer les processus naturels et la biologie des organismes existants (Remigi et al. 2019). L'évolution expérimentale a notamment récemment permis de mettre en évidence une nouvelle voie de régulation jusqu'à lors inconnue chez *R. solanacearum*, bactérie pourtant bien connue et étudiée depuis longtemps (Guidot et al. 2014; Capela et al. 2017). Un autre exemple du potentiel de l'EE dans la compréhension de l'évolution des bactéries est la découverte d'un mécanisme de co-transfert de gènes *nod* et d'hypermutagenèse chez les rhizobia qui a vraisemblablement favorisé leur évolution naturelle (Remigi et al. 2014).

Dans ce travail, à partir de deux observations liées à l'évolution expérimentale de *R. solanacearum* en symbiote de *Mimosa*, i. e. l'existence d'une délétion impactant l'infection intracellulaire et l'absence de capacité mutualiste des clones évolués de *Ralstonia*, nous avons identifié un nouveau gène *nod* impliqué dans la biosynthèse des FNs et démontré un contrôle spatio-temporel du mutualisme par la plante favorisant la dissémination du caractère fixateur dans des populations de rhizobia au cours de cycles de nodulation. Nous avons également abordé l'analyse cytologique des étapes précoces de l'infection de *M. pudica* par *R. solanacearum* portant des mutations adaptatives pour la symbiose, et par le rhizobium de référence *C. taiwanensis*.

V. 1. *noeM*, un nouveau gène de spécificité d'hôte

noeM est un nouveau gène *nod* de spécificité présent majoritairement chez des symbiotes de plantes du clade des mimosoidés, essentiellement des *Mimosa*. Nous avons démontré que ce gène est impliqué dans une structure atypique de FNs, dont le sucre réducteur est ouvert et oxydé. Nous avons montré que l'absence de *NoeM* affecte négativement la compétition de *C. taiwanensis* pour la nodulation de *M. pudica*. Néanmoins, au terme de ces travaux plusieurs questions restent en suspens.

Quelle est la fonction précise de NoeM ?

Nos collaborateurs chimistes de l'équipe de Sébastien Fort à Grenoble considèrent que la conversion des FNs classiques de *C. taiwanensis* (LCO V(C18:1, NMe, Cb, S)) en FNs atypiques (LCO open-oxidized (C18:1, NMe, Cb, S)) se fait en plusieurs étapes, nécessitant vraisemblablement plusieurs enzymes. Dans ce cas quelle serait l'étape catalysée par *NoeM* ?

Aucune activité de *NoeM* n'a été observée *in vitro* et *in vivo* en utilisant différents COs comme substrats (S. Fort, non publié), ce qui suggère que ces produits ne sont pas les substrats de *NoeM*. *NoeM* pourrait ne reconnaître que des COs ou LCOs avec des décorations spécifiques,

comme cela a été montré pour d'autres gènes *nod* (Broughton et al. 2000; Poinso et al. 2016). D'autres difficultés peuvent gêner la caractérisation de l'activité biochimique de cette protéine. Par exemple, dans les conditions d'étude *in vivo*, la présence de codons rares dans *noeM* peut être responsable d'une mauvaise traduction de l'ARN du gène *noeM* cloné chez *E. coli*. De plus, le fait que la protéine soit transmembranaire peut rendre difficile son expression chez *E. coli* et sa purification, empêchant sa caractérisation dans des conditions expérimentales *in vitro*. Néanmoins, d'autres protéines transmembranaires ont déjà pu être caractérisées, notamment la N-acetylglucosaminyltransferase NodC (John et al. 1988).

La caractérisation des FNs synthétisés par les souches rhizobiennes *Ensifer* sp. BR816 et *Rhizobium gallicum* bv. *gallicum* R602 avait révélé la présence de N-glucosamininitols (des FNs avec un sucre réducteur ouvert mais pas oxydé) en plus des N-glucosamines classiques (Snoeck et al. 2001; Soria-Díaz et al. 2006). Nos travaux ont permis d'établir la présence de *noeM* chez ces deux bactéries. Il est possible que lors des analyses de la structure des FNs de BR816 et R602, datant de 2001 et 2006 respectivement, des lipochitoooligosaccharides ouverts-oxydés n'aient pas été détectés (Snoeck et al. 2001; Soria-Díaz et al. 2006). Cela avait été le cas pour *C. taiwanensis* LMG19424 (Amadou et al. 2008). Si les structures décrites sont correctes et NoeM étant présente chez des souches produisant des N-glucosamininitols et des N-glucosamininitols oxydés, nous avons émis l'hypothèse que cette enzyme soit impliquée dans la transformation des N-glucosamines en N-glucosamininitols et qu'une autre enzyme, présente chez *C. taiwanensis* mais absente chez BR816 et R602 transformerait les N-glucosamininitols en N-glucosamininitols oxydés. De façon inattendue, le gène *noeM* de R602 n'a pas complété le mutant $\Delta noeM$ de *C. taiwanensis*. Il est possible que le gène de R602 ne soit pas bien exprimé dans le background de *C. taiwanensis*, une analyse d'expression de ces gènes par qRT-PCR pourrait nous permettre de contrôler cela. Il est aussi possible que le gène *noeM* de R602 ne soit pas fonctionnel comme c'est le cas pour d'autres souches de *Phaseolus vulgaris*. Dans ce cas, cela suggérerait que *noeM* soit plutôt impliqué dans la conversion de N-glucosamininitols en N-glucosamininitols oxydés. Néanmoins, si NoeM transformait les N-glucosamininitols en N-glucosamininitols oxydés, nous aurions dû retrouver des N-glucosamininitols avec le mutant $\Delta noeM$ de *C. taiwanensis*. Une nouvelle analyse des FNs synthétisés par BR816 et R602 pourrait permettre de s'assurer de l'absence de FNs ouverts-oxydés et fonder nos hypothèses de travail sur des bases plus solides.

Comme, il est peu vraisemblable qu'une seule enzyme soit incapable de réduire un N-glucosamine en N-glucosamininitol et ensuite d'oxyder ce dernier nous avons recherché de potentiels gènes candidats. L'organisation du génome autour de *noeM* ne montre pas d'autre gène en aval de la *nod*-box 2 chez LMG19424. Cependant, chez *B. phymatum* STM815 un autre gène, *BURPHP2_0639* codant pour une putative glycosyl hydrosylase, se trouve en aval de *noeM*. L'étude transcriptomique de STM815 a révélé que ce gène est fortement induit en réponse aux exsudats racinaires de *M. pudica* (Klonowska et al. 2018). Les glycosides hydrolases sont un large groupe d'enzymes qui catalyse les réactions d'hydrolyse, c'est-à-dire la rupture de liaisons osidiques entre deux ou plusieurs carbohydrates (glucides). On retrouve

chez *C. taiwanensis* LMG19424 un homologue de ce gène sur le pRalta, entouré de transposases, *pRalta_0250*, qui est annoté comme une endoglucanase. Cependant il n'apparaît pas surexprimé en présence d'exsudats racinaires de *M. pudica* (Klonowska et al. 2012). On retrouve des homologues de BURPHP2_0639 chez d'autres souches possédant *noeM* et qui sont des symbiotes de *Mimosa pudica* comme *B. phenoliruptrix* BR3459a (où la synténie est la même que chez STM815) mais aussi chez *R. mesoamericanum* STM3625. Par contre, on ne retrouve ce gène ni chez BR816 ni chez R602. La délétion du gène *pRalta_0250* chez *C. taiwanensis*, serait le meilleur moyen d'évaluer le rôle de ce gène dans la synthèse des FNs. Ainsi, si le gène est impliqué dans les étapes finales d'oxydation des FNs, un mutant $\Delta pRalta_0250$ devrait alors produire des FNs de type N-acetylglucosaminitol. Néanmoins, chez *C. taiwanensis*, l'absence de *nod*-box en amont de ce gène et le fait qu'il ne soit pas induit par les exsudats racinaires de *M. pudica* ne permettent pas de le décrire comme un gène *nod*.

Comment noeM s'est spécialisé dans la symbiose ?

L'analyse blastp de la séquence protéique de pRalta_0471 a permis d'identifier deux groupes bien distincts d'homologues. Le premier groupe avec une identité inférieure à 44% et le second groupe avec une identité supérieure à 70%. De façon intéressante toutes les protéines du second groupe appartiennent à des rhizobia. De plus, pour les souches dont le génome a été séquencé et assemblé, le gène codant est localisé en aval d'une *nod*-box sur le plasmide symbiotique. Nous avons émis l'hypothèse que les protéines appartenant au groupe 1 n'avaient pas la même fonction que celle du groupe 2, spécifique aux rhizobia. Malheureusement, nos travaux ne nous ont pas permis de conclure sur une éventuelle différence d'activité de NoeM en fonction du groupe d'appartenance des homologues. En effet, le remplacement de pRalta_0471 par le gène *noeM* de la souche *Paraburkholderia tuberum* WSM4176 n'a pas complété celui de *C. taiwanensis*, suggérant que cette protéine n'est pas active sur des FNs. Cependant, le contrôle positif (complémentation du mutant $\Delta noeM$ de *C. taiwanensis* par le gène *noeM* de R602) n'a pas fonctionné ne nous permettant pas de conclure sur le rôle fonctionnel de ces protéines. Ces travaux sont à répéter avec d'autres gènes.

Néanmoins, *noeM* est un gène *nod* impliqué dans la structure des FNs dont la spécialisation en gène *nod* pourrait s'être faite en plusieurs étapes. Il est probable que *noeM* provienne d'une duplication d'un gène codant pour une fatty acid hydroxylase. Les duplications géniques peuvent être à l'origine de multiples copies d'un même gène dans un génome, pouvant entraîner la possibilité de divergence fonctionnelle d'un gène (He and Zhang 2005). La divergence fonctionnelle d'un gène est le processus par lequel un gène change de fonction par rapport à celle qu'il possédait antérieurement. De nombreux exemples ont été décrits chez différents organismes dont *Arabidopsis* (Duarte et al. 2006), la drosophile (Assis and Bachtrog 2013) ainsi que des levures (DeLuna et al. 2008). L'acquisition d'une toute nouvelle capacité fonctionnelle, ou néofonctionnalisation, est le résultat d'un processus évolutif, impliquant une modification de l'activité enzymatique de la protéine et/ou une modification des conditions de son expression (Rastogi and Liberles 2005; Conant and Wolfe 2008). La

spécialisation de *noeM* dans la modification de FNs, pourrait résulter uniquement de son raccordement au réseau de régulation de la nodulation. En effet, chez tous les rhizobia séquencés portant *noeM*, ce gène est localisé sur un plasmide symbiotique en aval d'une *nod*-box. Or, certaines souches (par exemple *Burkholderia* sp. JPY251) possèdent deux copies de *noeM* avec seulement l'une de ces deux copies sous contrôle de la *nod*-box, suggérant une duplication du gène en aval de la *nod*-box et donc néofonctionnalisation par raccordement au réseau *nod*. La plupart des gènes *nod*, notamment *nodB* et *nodC*, sont supposés être issus d'une divergence fonctionnelle (Hirsch et al. 2001).

Comment expliquer la sélection pour $\Delta noeM$ dans l'expérience d'évolution ?

Dans les populations naturelles bactériennes les mutations ponctuelles, délétions et réarrangements génomiques apparaissent à une fréquence de 10^{-11} (Wielgoss et al. 2011). Ces mutations peuvent être neutres, délétères ou bénéfiques. En fonction de leur impact elles sont éliminées ou conservées (sélection purificatrice ou positive). Dans l'expérience d'évolution, la délétion de *noeM* a un effet négatif à la fois sur l'infection et la nodulation, son maintien est donc surprenant. Ce maintien suggère donc que la délétion de *noeM* est physiquement liée à une mutation adaptative. Une autre possibilité est que la perte de *noeM* a un autre effet adaptatif plus fort que son effet délétère dans les conditions de l'expérience.

La première hypothèse est que la remontée en fréquence de la délétion de *noeM* pourrait être en lien avec une autre mutation fortement adaptative. C'est un phénomène que l'on appelle l'auto-stop génétique (ou *genetic hitchhiking*) (Barton 2000). Par exemple, si une mutation fortement adaptative (pour la croissance, la nodulation, etc...) survient en même temps dans le génome que la délétion de *noeM*, ces deux mutations vont remonter en fréquence dans la population. D'autant plus que nous sommes dans un contexte d'hypermutable, encore plus marqué dans les lignées longues, favorisant le hitchhiking. Afin d'identifier la ou les mutations liées par auto-stop génétique avec la délétion de *noeM*, il faudrait mesurer l'impact sur le fitness *in planta* de chacune des mutations qui se sont fixées en même temps que la délétion de *noeM*.

Notre seconde hypothèse est que la délétion de *noeM* soit sélectionnée dans l'expérience d'évolution en lien avec la longueur des cycles. En effet, cette délétion favorise un retard de nodulation dans les lignées longues. Or, les travaux menés dans l'équipe avant mon arrivée ont montré que la complémentation de *noeM* dans le clone évolué R16 restaure non seulement l'infection intracellulaire mais améliore également la nodulation qui était très décalée. On peut penser que dans les lignées longues le retard à la nodulation est une stratégie pour conserver un meilleur fitness étant donné que les souches ont une mauvaise persistance dans les cellules du nodule. Ainsi, la nodulation étant retardée, l'infection des nodules est aussi retardée et donc la mortalité des bactéries est retardée à 42 dpi, timing de la récolte des nodules. Le fait qu'au cours des cycles, la capacité à noduler est de plus en plus retardé semble conforter cette idée.

V. 2. La dissémination du caractère fixateur d'azote

Une grande question qui intéresse les scientifiques depuis longtemps est de comprendre ce qui permet le maintien du mutualisme. La théorie veut que, au sein des populations bactériennes symbiotiques évoluent des « cheaters » qui perdent la capacité à fixer l'azote mais profitent de la niche nodule avec à terme un risque de destruction du mutualisme (Heath and Tiffin 2009; Jones et al. 2015). Afin d'éviter de mettre à mal la stabilité de la symbiose mutualiste, les plantes auraient mis en place des sanctions pour éviter la propagation des cheaters (Kiers et al. 2003; Oono et al. 2011; Fujita et al. 2014). La plupart des études se sont intéressées aux mécanismes qui permettraient le maintien du mutualisme au cours de l'évolution (Kiers et al. 2003; Regus et al. 2017) alors que peu d'études se sont intéressées à l'aspect contraire, à savoir l'expansion du mutualiste. En effet, il est vraisemblable qu'au cours de l'évolution et dans certains cas –par exemple lors d'un transfert longue distance- le protorhizobium obtenu après acquisition des gènes *nod-nif* ne soit pas fixateur d'azote (Nandasena et al. 2006). Dans ce cas, dans quelles conditions des mutants Fix+ -si ils apparaissent- vont-ils envahir la population et le mutualisme peut-il se propager ? Cette question nous a amenée à étudier le destin d'une population mixte de souches isogéniques Fix+/Fix- de *C. taiwanensis* au cours de l'infection de *M. pudica*. Il est à souligner, que le terme Fix- est intégratif puisqu'un rhizobium peut être Fix- pour différentes raisons (incapable de noduler, d'infecter, de persister ou bien de fixer l'azote), mais dans nos travaux nous sommes intéressés uniquement à la fixation d'azote *per se*.

Un contrôle au niveau cellulaire

Il avait déjà été démontré que dans certains systèmes les Fix- dégénèrent en simple inoculation (Hirsch and Smith 1987; Berrabah et al. 2015), mais pas dans toutes les interactions rhizobium-légumineuse (Marco et al. 2009; Ling et al. 2013). De nombreuses études avaient montré une dégénérescence spécifique des nodules non fixateurs suite à une co-inoculation avec des rhizobia Fix+ et Fix-, mais dans la plupart des cas il s'agissait de souches naturelles dont l'origine du caractère non fixateur n'avait pas été analysé (Chen and Thornton 1940; Singleton and Stockinger 1983; Simms et al. 2006). Grâce à une étude cinétique du devenir de bactéries isogéniques uniquement différenciées par leur caractère fixateur d'azote *per se*, nous avons pu montrer une mortalité spécifique des bactéroïdes Fix- et déterminer le moment à partir duquel la dégénérescence précoce des bactéries non fixatrices se met en place. De plus, notre travail a permis de mettre en évidence que les « sanctions végétales » agissent au niveau cellulaire. L'existence de « cell-autonomous sanctions » a également été démontré au même moment par les travaux de Regus et de ses collaborateurs qui ont obtenus des résultats similaires sur un autre système, avec *Lotus japonicus* dont les nodules sont déterminés (Regus et al. 2017). Dans notre travail nous avons utilisé des phénotypes extrêmes, à savoir des bactérie Fix+ et Fix-. La question reste de savoir à partir de quel seuil de fixation d'azote, si il y en a un, les sanctions sont levées. Des travaux ont déjà été initiés par Kiers et ses collaborateurs afin de savoir si les sanctions interviennent

à des niveaux intermédiaires de fixation d'azote. Ils ont ainsi mis en place un système artificiel où la concentration de N₂ était modulée à différents niveaux intermédiaires afin de manipuler la fixation des rhizobia inoculés (Kiers et al. 2006). Ils ont alors démontré que le niveau de fixation et le fitness des rhizobia était significativement corrélés suggérant que la force des sanctions variait en fonction du pourcentage de N₂ et donc du niveau de fixation (Kiers et al. 2006).

Un mécanisme de sanction végétal?

On sait depuis longtemps que la plante est sensible au statut azoté puisqu'elle est capable de réguler la formation de nodules en fonction de ses besoins en azote (Reid et al. 2011). La perte de viabilité des Fix- dans notre système conforte l'hypothèse de sanctions végétales (Kiers et al. 2003; Oono et al. 2011). Cette hypothèse, bien que non formellement démontrée est la plus plausible. Il n'est cependant pas impossible que l'inactivation de la nitrogénase engendre l'accumulation d'électrons –ou d'autres phénomènes- qui pourrait être létale pour les bactéroïdes.

A ce jour les mécanismes moléculaires impliqués dans la mise en place de sanctions sont inconnus. Dans ses travaux, Kiers émet l'hypothèse que c'est la plante qui sanctionne les bactéroïdes les moins efficaces en leur supprimant l'apport de O₂ dont ils ont besoin (Kiers et al. 2003). L'existence de mutants de plante qui laisseraient viables des non fixateurs pourrait permettre d'analyser les mécanismes des sanctions. Les nombreuses études menées sur les plantes modèle *M. truncatula* (Krishnakumar et al. 2015) et *L. japonicus* (Perry et al. 2003; Sato et al. 2008) sont à l'origine de nombreux mutants. Néanmoins, à notre connaissance aucun mutant de plante qui ne tue pas les Fix- n'a encore été mis en évidence, et le criblage de tels mutants serait long et difficile.

Une alternative plus raisonnable serait de construire une gamme de mutants bactériens isogéniques qui ne se différencieraient que par leur capacité à fixer l'azote plus ou moins importante. L'utilisation de couples de bactéries ayant des capacités fixatrices intermédiaires pourrait permettre de définir si la mortalité des bactéroïdes intervient à partir d'un seuil de fixation minimal (sanction intrinsèque), ou si elle dépend des couples utilisés (sanction relative), c'est-à-dire si une bactérie est sanctionnée non pas uniquement en fonction de sa capacité fixatrice mais également en fonction de la présence de fixation d'azote issue d'autres symbiotes et donc du statut azoté de la plante. Dans ce cas on pourrait affirmer l'hypothèse sanction.

Évolution naturelle et expérimentale du mutualisme

Dans le contexte du laboratoire nos travaux avec des souches isogéniques ont permis de montrer que le caractère Fix+ est sélectionné et que les sanctions végétales permettent la dissémination du caractère au cours de cycles de nodulation. La question est de savoir dans quelle mesure ce processus opère dans la nature. La fixation d'azote étant un des caractères symbiotiques il peut être contrebalancé par d'autres caractères, plus fortement adaptatifs.

Ainsi il a été montré que certains plasmides de *S. meliloti* augmentent la compétitivité pour la nodulation mais restreignent la capacité de fixation de l'azote (Crook et al. 2012). Il s'avère que les plasmides sont porteurs d'une « pHR-encoded metallopeptidase » (HrrP), une peptidase capable d'empêcher l'établissement des étapes tardives de l'infection (Price et al. 2015). En effet, la peptidase en question serait capable de cliver les NCR de *Medicago*, ce qui déclencherait une dégénération prématurée des bactéroïdes, en lien avec l'absence de fixation qu'ils ont observé dans leurs précédents travaux (Crook et al. 2012; Price et al. 2015). Ces travaux démontrent que dans la nature le transfert horizontal de gènes peut convertir un rhizobium qui était fixateur en symbiote plus compétitifs mais moins efficient (Crook et al. 2012). Par ailleurs dans la nature les populations de rhizobia ne sont pas homogènes et sont en compétition avec d'autres populations de rhizobia qui peuvent interférer avec ce processus. Ces observations nous rappellent que les travaux menés en laboratoire ne reflètent pas la complexité des phénomènes évolutifs que l'on peut retrouver dans la nature.

Néanmoins, nos résultats ont permis de déterminer les conditions favorables à l'émergence de clones fixateurs en cas d'apparition d'un allèle Fix+ au sein d'une population majoritairement Fix-. Ainsi, nous avons conclu que l'utilisation de cycles de 28 jours dans l'expérience d'évolution permettraient la remontée en fréquence d'une sous population fixatrice, si elle émerge. C'est la durée qui sera utilisée lorsque les clones évolués auront acquis une persistance équivalente à celle d'un mutant *nifH* dans notre expérience d'évolution.

V.3. Les étapes tardives de l'infection du nodules sont-elles influencées dès les premières étapes de l'infection du poil absorbant ?

L'étude développementale des nodules de *M. truncatula* et de certains de ses mutants ont permis de déterminer que l'organogenèse des nodules est dépendante d'évènements spatio-temporels précis, comme le développement des cordons d'infection qui doit avoir lieu avant la division péricline des cellules du méristème (Xiao et al. 2014). Cependant, cette étude ne portent que sur des mutants de plantes, sans prendre en compte le côté bactérien. Il est possible que des mutants bactériens aient le même effet sur le développement du nodule. L'expérience d'évolution a permis l'obtention d'un matériel génétique unique en son genre, avec une collection de mutants bactériens quasi-isogéniques dont les propriétés symbiotiques sont variables (Marchetti et al. 2016). L'étude préliminaire menée sur l'infection racinaire de *M. pudica* par des mutants de *R. solanacearum* semble suggérer un retard de l'initiation de l'infection des clones de *R. solanacearum* comparé au rhizobium de référence, *C. taiwanensis*. A ce stade, il paraît évident de terminer les répétitions techniques des clones de *R. solanacearum* qui sont en cours et d'ajouter les clones portant les mutations adaptives pour l'amélioration de l'infection des cellules du nodule comme *hrpGQ81* et *phcBR22C* ou *efpRE66K* dans un background de *hrpGQ81*. Notre hypothèse est que ces clones portant des mutations améliorant la compétition pour la nodulation et améliorant l'infection intracellulaire, devrait

avoir un délai d'apparition des sites d'infection, cordons d'infection et nodules réduit comparé aux souches GMI1000 pRalta et GMI1000 pRalta *hrcV*. Cela impliquerait que l'invasion des bactéries dans les cellules de l'hôte soit dépendante de l'activation des gènes de développement et/ou de l'inactivation des défenses végétales lors de la progression des bactéries dans les cordons d'infection. La corrélation faite entre l'amélioration de l'infection intracellulaire et l'amélioration de la compétition vont dans ce sens (Guan et al. 2013). L'analyse transcriptomique de racines *M. pudica* infectées par différentes souches bactériennes en fonction de leur capacité de nodulation et d'infection aux étapes précoces de l'infection devrait permettre l'identification de gènes impactés lors de l'infection par les différentes souches bactériennes et dont l'expression est modulée au cours des transitions phénotypiques bactériennes.

Références bibliographiques :

- Akkermans, A. D. L., Abdulkadir, S., and Trinick, M. J. 1978. N₂-fixing root nodules in Ulmaceae: *Parasponia* or (and) *Trema* spp.? *Plant Soil*. 49:711–715
- Aldon, D., Brito, B., Boucher, C., Genin, S., and Boucher, C. A. 2000. A bacterial sensor of plant cell contact controls the transcriptional induction of *Ralstonia solanacearum* pathogenicity genes. *EMBO J*. 19:2304–2314
- Alexander, C., and Rietschel, E. T. 2001. Invited review: Bacterial lipopolysaccharides and innate immunity. *J. Endotoxin Res*. 7:167–202
- Amadou, C., Mangenot, S., Glew, M., Bontemps, C., Capela, D., Dossat, C., Marchetti, M., Servin, B., Saad, M., Schenowitz, C., Batut, J., and Masson-boivin, C. 2008. Genome sequence of the β -rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Res*. :1472–1483
- Amor, B. Ben, Shaw, S. L., Oldroyd, G. E. D., Maillet, F., Penmetsa, R. V., Cook, D., Long, S. R., Denarie, J., and Gough, C. 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
- Aoki, S., Ito, M., and Iwasaki, W. 2013. From β - to α -proteobacteria: the origin and evolution of rhizobial nodulation genes nodJ. *Mol. Biol. Evol*. 30:2494–508
- Appleby, C. A. 1984. Leghemoglobin and Rhizobium Respiration. *Plant Physiol*.
- Arlat, M., Gough, C. L., Zischek, C., Barberis, P. A., Trigalet, A., and Boucher, C. A. 1992. Transcriptional organization and expression of the large *hrp* gene cluster of *Pseudomonas solanacearum*. *Mol. Plant. Microbe. Interact*. 5:187–93
- Assis, R., and Bachtrog, D. 2013. Neofunctionalization of young duplicate genes in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A*. 110:17409–14
- Azani, N., Babineau, M., Bailey, C. D., Banks, H., Barbosa, A., Pinto, R. B., Boatwright, J., Borges, L., Brown, G., Bruneau, A., Candido, E., Cardoso, D., Chung, K.-F., Clark, R., Conceição, A. de S., Crisp, M., Cubas, P., Delgado-Salinas, A., Dexter, K., Doyle, J., Duminil, J., Egan, A., De La Estrella, M., Falcão, M., Filatov, D., Fortuna-Perez, A. P., Fortunato, R., Gagnon, E., Gasson, P., Rando, J. G., Azevedo Tozzi, A. M. G. de, Gunn, B., Harris, D., Haston, E., Hawkins, J., Herendeen, P., Hughes, C., Iganci, J. V., Javadi, F., Kanu, S. A., Kazempour-Osaloo, S., Kite, G., Klitgaard, B., Kochanowski, F., Koenen, E. M., Kovar, L., Lavin, M., Roux, M. le, Lewis, G., de Lima, H., López-Roberts, M. C., Mackinder, B., Maia, V. H., Malécot, V., Mansano, V., Marazzi, B., Mattapha, S., Miller, J., Mitsuyuki, C., Moura, T., Murphy, D., Nageswara-Rao, M., Nevado, B., Neves, D., Ojeda, D., Pennington, R. T., Prado, D., Prenner, G., de Queiroz, L. P., Ramos, G., Ranzato Filardi, F., Ribeiro, P., Rico-Arce, M. de L., Sanderson, M., Santos-Silva, J., São-Mateus, W. B., Silva, M. S., Simon, M., Sinou, C., Snak, C., de Souza, É., Sprent, J., Steele, K., Steier, J., Steeves, R., Stirton, C., Tagane, S., Torke, B., Toyama, H., Cruz, D. T. da, Vatanparast, M., Wieringa, J., Wink, M., Wojciechowski, M., Yahara, T., Yi, T., and Zimmerman, E. 2017. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny – The Legume Phylogeny Working Group (LPWG). *Taxon*. 66:44–77
- Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E., and Kim, J. F. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature*. 461:1243–1247
- Barton, N. H. 2000. Genetic hitchhiking B. Charlesworth and P.H. Harvey, eds. *Philos. Trans. R. Soc. London. Ser. B Biol. Sci*. 355:1553–1562
- Bartsev, A. V., Deakin, W. J., Boukli, N. M., McAlvin, C. B., Stacey, G., Malnoë, P., Broughton, W. J., and Staehelin, C. 2004. NopL, an effector protein of *Rhizobium* sp. NGR234, thwarts activation of plant defense reactions. *Plant Physiol*. 134:871–9
- Becquart-de Kozak, I., Reuhs, B. L., Buffard, D., Breda, C., Kim, J. S., Esnault, R., and Kondorosi, A. 1997. Role of the K-antigen subgroup of capsular polysaccharides in the early recognition

- process between *Rhizobium meliloti* and alfalfa leaves. Pages 114–123 in: *Molecular Plant-Microbe Interactions*, The American Phytopathological Society.
- Berrabah, F., Ratet, P., and Gourion, B. 2015. Multiple steps control immunity during the intracellular accommodation of rhizobia. *J. Exp. Bot.* 66:1977–85
- Bhagwat, A. A., Mithöfer, A., Pfeffer, P. E., Kraus, C., Spickers, N., Hotchkiss, A., Ebel, J., and Keister, D. L. Further Studies of the Role of Cyclic β -Glucans in Symbiosis. An *ndvC* Mutant of *Bradyrhizobium japonicum* Synthesizes Cyclodecakis-(1 \rightarrow 3)- β -Glucosyl. *Plant Physiol.* 119:1057–1064
- Blair, M. W., Pedraza, F., Buendia, H. F., Gaitán-Solís, E., Beebe, S. E., Gepts, P., and Tohme, J. 2003. Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 107:1362–1374
- Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., Goormachtig, S., and Giraud, E. 2011. Nodulation of *Aeschynomene afraspera* and *A. indica* by Photosynthetic *Bradyrhizobium* Sp. Strain ORS285: The Nod-Dependent Versus the Nod-Independent Symbiotic Interaction. *Mol. Plant-Microbe Interact.* 24:1359–1371
- Bontemps, C., Elliott, G. N., Simon, M. F., Dos Reis Júnior, F. B., Gross, E., Lawton, R. C., Neto, N. E., de Fátima Loureiro, M., De Faria, S. M., Sprent, J. I., James, E. K., and Young, J. P. W. 2010. Burkholderia species are ancient symbionts of legumes. *Mol. Ecol.* 19:44–52
- Boucher, C. A., Barberis, P. A., and Demery, D. A. 1985. Transposon Mutagenesis of *Pseudomonas solanacearum*: Isolation of Tn5-Induced Avirulent Mutants. *Microbiology.* 131:2449–2457
- Breedveld, M. W., Cremers, H. C., Batley, M., Posthumus, M. A., Zevenhuizen, L. P., Wijffelman, C. A., and Zehnder, A. J. 1993. Polysaccharide synthesis in relation to nodulation behavior of *Rhizobium leguminosarum*. *J. Bacteriol.* 175:750–7
- Breedveld, M. W., and Miller, K. J. 1994. Cyclic beta-glucans of members of the family Rhizobiaceae. *Microbiol. Mol. Biol. Rev.* 58
- Brito, B., Marena, M., Barberis, P., Boucher, C., and Genin, S. 1999. *prhJ* and *hrpG*, two new components of the plant signal-dependent regulatory cascade controlled by PrhA in *Ralstonia solanacearum*. *Mol. Microbiol.* 31:237–251
- Brockhurst, M. A., Colegrave, N., and Rozen, D. E. 2011. Next-generation sequencing as a tool to study microbial evolution. *Mol. Ecol.* 20:972–980
- Broughton, W. J., Jabbouri, S., and Perret, X. 2000. Keys to symbiotic harmony. *J. Bacteriol.* 182:5641–52
- Broughton, W. J., and Perret, X. 1999. Genealogy of legume-Rhizobium symbioses. *Curr. Opin. Plant Biol.* 2:305–311
- Brown, D. B., Huang, Y.-C., Kannenberg, E. L., Sherrier, D. J., and Carlson, R. W. 2011. An *acpXL* Mutant of *Rhizobium leguminosarum* bv. *phaseoli* Lacks 27-Hydroxyoctacosanoic Acid in Its Lipid A and Is Developmentally Delayed during Symbiotic Infection of the Determinate Nodulating Host Plant *Phaseolus vulgaris*. *J. Bacteriol.* 193:4766–4778
- Buckling, A., Craig Maclean, R., Brockhurst, M. A., and Colegrave, N. 2009. The Beagle in a bottle. *Nature.* 457:824–829
- Burgess, B. K., and Lowe, D. J. 1996. Mechanism of Molybdenum Nitrogenase.
- Capela, D., Marchetti, M., Clérissi, C., Perrier, A., Guetta, D., Gris, C., Valls, M., Jauneau, A., Cruveiller, S., Rocha, E. P. C., and Masson-Boivin, C. 2017. Recruitment of a Lineage-Specific Virulence Regulatory Pathway Promotes Intracellular Infection by a Plant Pathogen Experimentally Evolved into a Legume Symbiont. *Mol. Biol. Evol.* 34:2503–2521
- Capoen, W., Oldroyd, G., Goormachtig, S., and Holsters, M. 2010. *Sesbania rostrata*: a case study of natural variation in legume nodulation. *New Phytol.* 186:340–345
- Capoen, W., Sun, J., Wysham, D., Otegui, M. S., Venkateshwaran, M., Hirsch, S., Miwa, H., Downie, J. A., Morris, R. J., Ané, J. M., and Oldroyd, G. E. D. 2011. Nuclear membranes control symbiotic calcium signaling of legumes. *Proc. Natl. Acad. Sci. U. S. A.* 108:14348–14353
- Carlson, R. W., Price, N. P., and Stacey, G. 1994. The biosynthesis of rhizobial lipo-oligosaccharide nodulation signal molecules. *Mol. Plant. Microbe. Interact.* 7:684–95

- Chandler, M. R. 1978. Some Observations on Infection of *Arachis hypogaea* L. by *Rhizobium*. *J. Exp. Bot.* 29:749–755
- Charpentier, M., Sun, J., Martins, T. V., Radhakrishnan, G. V., Findlay, K., Soumpourou, E., Thouin, J., Véry, A. A., Sanders, D., Morris, R. J., and Oldroyd, G. E. D. 2016. Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science* (80-.). 352:1102–1105
- Chen, H. K., and Thornton, H. G. 1940. The structure of 'ineffective' nodules and its influence on nitrogen fixation. *Proc. R. Soc. London. Ser. B - Biol. Sci.* 129:208–229
- Chen, W.-M., James, E. K., Chou, J.-H., Sheu, S.-Y., Yang, S.-Z., and Sprent, J. I. 2005. beta-Rhizobia from *Mimosa pigra*, a newly discovered invasive plant in Taiwan. *New Phytol.* 168:661–675
- Chen, W.-M., James, E. K., Prescott, A. R., Kierans, M., and Sprent, J. I. 2003. Nodulation of *Mimosa* spp. by the beta-proteobacterium *Ralstonia taiwanensis*. *Mol. Plant. Microbe. Interact.* 16:1051–61
- Chen, W. M., Laevens, S., Lee, T. M., Coenye, T., De Vos, P., Mergeay, M., and Vandamme, P. 2001. *Ralstonia taiwanensis* sp. nov., isolated from root nodules of *Mimosa* species and sputum of a cystic fibrosis patient. *Int. J. Syst. Evol. Microbiol.* 51:1729–1735
- Cheng, H. P., and Walker, G. C. 1998. Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. *J. Bacteriol.* 180:5183–91
- Clerissi, C., Touchon, M., Capela, D., Tang, M., Cruveiller, S., Parker, M. A., Moulin, L., Masson-Boivin, C., and Rocha, E. P. C. 2018. Parallels between experimental and natural evolution of legume symbionts. *Nat. Commun.* 9:2264
- Clough, S. J., Lee, K. E., Schell, M. A., and Denny, T. P. 1997. A two-component system in *Ralstonia* (*Pseudomonas*) *solanacearum* modulates production of PhcA-regulated virulence factors in response to 3-hydroxypalmitic acid methyl ester. *J. Bacteriol.* 179:3639–48
- Coll, N. S., and Valls, M. 2013. Current knowledge on the *Ralstonia solanacearum* type III secretion system. *Microb. Biotechnol.* 6:n/a-n/a
- Conant, G. C., and Wolfe, K. H. 2008. Turning a hobby into a job: How duplicated genes find new functions. *Nat. Rev. Genet.* 9:938–950
- Cook, D. R. 1999. *Medicago truncatula*--a model in the making! *Curr. Opin. Plant Biol.* 2:301–4
- Cooper, J. E. 2007. Early interactions between legumes and rhizobia: disclosing complexity in a molecular dialogue. *J. Appl. Microbiol.* 103:1355–1365
- Crespo-Rivas, J. C., Margaret, I., Hidalgo, Á., Buendía-Clavería, A. M., Ollero, F. J., López-Baena, F. J., del Socorro Murdoch, P., Rodríguez-Carvajal, M. A., Soria-Díaz, M. E., Reguera, M., Lloret, J., Sumpton, D. P., Mosely, J. A., Thomas-Oates, J. E., van Brussel, A. A. N., Gil-Serrano, A., Vinardell, J. M., and Ruiz-Sainz, J. E. 2009. *Sinorhizobium fredii* HH103 *cgs* Mutants Are Unable to Nodulate Determinate- and Indeterminate Nodule-Forming Legumes and Overproduce an Altered EPS. *Mol. Plant-Microbe Interact.* 22:575–588
- Crook, M. B., Lindsay, D. P., Biggs, M. B., Bentley, J. S., Price, J. C., Clement, S. C., Clement, M. J., Long, S. R., and Griffiths, J. S. 2012. Rhizobial Plasmids That Cause Impaired Symbiotic Nitrogen Fixation and Enhanced Host Invasion. *Mol. Plant-Microbe Interact.* 25:1026–1033
- Cruz, A. P. Z., Ferreira, V., Pianzola, M. J., Siri, M. I., Coll, N. S., and Valls, M. 2014. A Novel, Sensitive Method to Evaluate Potato Germplasm for Bacterial Wilt Resistance Using a Luminescent *Ralstonia solanacearum* Reporter Strain. *Mol. Plant-Microbe Interact.* 27:277–285
- Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., Pierre, O., Fardoux, J., Chaintreuil, C., Nguyen, P., Gressent, F., Da Silva, C., Poulain, J., Wincker, P., Rofidal, V., Hem, S., Barrière, Q., Arrighi, J.-F., Mergaert, P., and Giraud, E. 2015. Convergent Evolution of Endosymbiont Differentiation in Dalbergioid and Inverted Repeat-Lacking Clade Legumes Mediated by Nodule-Specific Cysteine-Rich Peptides. *Plant Physiol.* 169:1254–65
- D’Haeze, W., and Holsters, M. 2002. Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology.* 12:79R-105R
- Dallinger, W. 1878. On the life-history of a minute septic organisms: with an account of experiments made to determine its thermal death point. *Proc. R. Soc. London.* 27:332–350

- Daubech, B., Poinso, V., Klonowska, A., Capela, D., Chaintreuil, C., Moulin, L., Marchetti, M., and Masson-Boivin, C. 2019. noeM, a New Nodulation Gene Involved in the Biosynthesis of Nod Factors with an Open-Chain Oxidized Terminal Residue and in the Symbiosis with *Mimosa pudica*. *Mol. Plant-Microbe Interact.* 32:1635–1648
- Daubech, B., Remigi, P., Doin de Moura, G., Marchetti, M., Pouzet, C., Auriac, M.-C., Gokhale, C. S., Masson-Boivin, C., and Capela, D. 2017. Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation. *Elife.* 6:e28683
- Dawson, G. A. 1980. Nitrogen fixation by lightning. *J. Atmos. Sci.* 37:174–178
- Deardorff, E. R., Fitzpatrick, K. A., Jerzak, G. V. S., Shi, P.-Y., Kramer, L. D., and Ebel, G. D. 2011. West Nile Virus Experimental Evolution in vivo and the Trade-off Hypothesis C.O. Wilke, ed. *PLoS Pathog.* 7:e1002335
- Debellé, F., Rosenberg, C., and Dénarié, J. 1992. The Rhizobium, Bradyrhizobium, and Azorhizobium NodC proteins are homologous to yeast chitin synthases. *Mol. Plant. Microbe. Interact.* 5:443–6
- Delaspre, F., Nieto Peñalver, C. G., Saurel, O., Kiefer, P., Gras, E., Milon, A., Boucher, C., Genin, S., and Vorholt, J. A. 2007. The *Ralstonia solanacearum* pathogenicity regulator HrpB induces 3-hydroxy-oxindole synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 104:15870–5
- DeLuna, A., Vetsigian, K., Shore, N., Hegreness, M., Colón-González, M., Chao, S., and Kishony, R. 2008. Exposing the fitness contribution of duplicated genes. *Nat. Genet.* 40:676–681
- Denny, T. P. 1991. Genetic Evidence that Extracellular Polysaccharide Is a Virulence Factor of *Pseudomonas solanacearum*. *Mol. Plant-Microbe Interact.* 4:198
- Dixon, R., and Kahn, D. 2004. Genetic regulation of biological nitrogen fixation. *Nat. Rev. Microbiol.* 2:621–631
- Doyle, J. J. 2011. Phylogenetic Perspectives on the Origins of Nodulation. *Mol. Plant-Microbe Interact.* 24:1289–1295
- Dreyfus, B., Garcia, J. L., and Gillis, M. 1988. Characterization of *Azorhizobium caulinodans* gen. nov., sp. nov., a Stem-Nodulating Nitrogen-Fixing Bacterium Isolated from *Sesbania rostrata*. *Int. J. Syst. Bacteriol.* 38:89–98
- Duarte, J. M., Cui, L., Wall, P. K., Zhang, Q., Zhang, X., Leebens-Mack, J., Ma, H., Altman, N., and dePamphilis, C. W. 2006. Expression Pattern Shifts Following Duplication Indicative of Subfunctionalization and Neofunctionalization in Regulatory Genes of *Arabidopsis*. *Mol. Biol. Evol.* 23:469–478
- Elena, S. F., Agudelo-Romero, P., Carrasco, P., Codoñer, F. M., Martín, S., Torres-Barceló, C., and Sanjuán, R. 2008. Experimental evolution of plant RNA viruses. *Heredity (Edinb).* 100:478–483
- Elena, S. F., and Lenski, R. E. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* 4:457–469
- van Elsas, J. D., Kastelein, P., van Bekkum, P., van der Wolf, J. M., de Vries, P. M., and van Overbeek, L. S. 2000. Survival of *Ralstonia solanacearum* Biovar 2, the Causative Agent of Potato Brown Rot, in Field and Microcosm Soils in Temperate Climates. *Phytopathology.* 90:1358–1366
- Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kaló, P., and Kiss, G. B. 2002. A receptor kinase gene regulating symbiotic nodule development. *Nature.* 417:962–966
- Erismann, J. W., Sutton, M. A., Galloway, J., Klimont, Z., and Winiwarter, W. 2008. How a century of ammonia synthesis changed the world. *Nat. Geosci.* 2008 110.
- Estrada-de los Santos, P., Martínez-Aguilar, L., López-Lara, I. M., and Caballero-Mellado, J. 2012. *Cupriavidus alkaliphilus* sp. nov., a new species associated with agricultural plants that grow in alkaline soils. *Syst. Appl. Microbiol.* 35:310–314
- Falcone Ferreyra, M. L., Rius, S. P., and Casati, P. 2012. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* 3:222
- De Faria, S. M., Hay, G. T., and Sprent, J. I. 1988. Entry of Rhizobia into Roots of *Mimosa scabrella* Benth. Occurs between Epidermal Cells. *Microbiology.* 134:2291–2296
- Fauvert, M., and Michiels, J. 2008. Rhizobial secreted proteins as determinants of host specificity in the rhizobium-legume symbiosis. *FEMS Microbiol. Lett.* 285:1–9
- Fellay, R., Hanin, M., Montorzi, G., Frey, J., Freiberg, C., Golinowski, W., Staehelin, C., Broughton, W.

- J., and Jabbouri, S. 1998. *nodD 2* of *Rhizobium* sp. NGR234 is involved in the repression of the *nodABC* operon. *Mol. Microbiol.* 27:1039–1050
- Firmin, J. L., Wilson, K. E., Carlson, R. W., Davies, A. E., and Downie, J. A. 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.* 10:351–360
- Florentino, L. A., Jaramillo, P. M. D., Silva, K. B., Silva, J. S. da, Oliveira, S. M. de, and Moreira, F. M. de S. 2012. Physiological and symbiotic diversity of *Cupriavidus necator* strains isolated from nodules of Leguminosae species. *Sci. Agric.* 69:247–258
- Forsberg, L. S., and Reuhs, B. L. 1997. Structural characterization of the K antigens from *Rhizobium fredii* USDA257: evidence for a common structural motif, with strain-specific variation, in the capsular polysaccharides of *Rhizobium* spp. *J. Bacteriol.* 179:5366–71
- Fournier, J., Teillet, A., Chabaud, M., Ivanov, S., Genre, A., Limpens, E., de Carvalho-Niebel, F., and Barker, D. G. 2015. Remodeling of the infection chamber before infection thread formation reveals a two-step mechanism for rhizobial entry into the host legume root hair. *Plant Physiol.* 167:1233–42
- Fournier, J., Timmers, A. C. J., Sieberer, B. J., Jauneau, A., Chabaud, M., and Barker, D. G. 2008. Mechanism of infection thread elongation in root hairs of *Medicago truncatula* and dynamic interplay with associated rhizobial colonization. *Plant Physiol.* 148:1985–95
- Frederickson, M. E. 2013. Rethinking mutualism stability: Cheaters and the evolution of sanctions. *Q. Rev. Biol.* 88:269–295
- French, J. R. J., Turner, G. L., and Bradbury, J. F. 1976. Nitrogen Fixation by Bacteria from the Hindgut of Termites. *J. Gen. Microbiol.* 95:202–206
- Fujita, H., Aoki, S., and Kawaguchi, M. 2014. Evolutionary dynamics of nitrogen fixation in the legume-rhizobia symbiosis. *PLoS One.* 9:e93670
- Gaby, J. C., Rishishwar, L., Valderrama-Aguirre, L. C., Green, S. J., Valderrama-Aguirre, A., Jordan, I. K., and Kostka, J. E. 2018. Diazotroph Community Characterization via a High-Throughput nifH Amplicon Sequencing and Analysis Pipeline. *Appl. Environ. Microbiol.* 84:e01512-17
- Gage, D. J. 2002. Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*. *J. Bacteriol.* 184:7042–6
- Gage, D. J. 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol. Mol. Biol. Rev.* 68:280–300
- Gehlot, H. S., Tak, N., Kaushik, M., Mitra, S., Chen, W.-M., Poweleit, N., Panwar, D., Poonar, N., Parihar, R., Tak, A., Sankhla, I. S., Ojha, A., Rao, S. R., Simon, M. F., Reis Junior, F. B. dos, Perigolo, N., Tripathi, A. K., Sprent, J. I., Young, J. P. W., James, E. K., and Gyaneshwar, P. 2013. An invasive *Mimosa* in India does not adopt the symbionts of its native relatives. *Ann. Bot.* 112:179–196
- Genin, S. 2010. Molecular traits controlling host range and adaptation to plants in *Ralstonia solanacearum*. *New Phytol.* 187:920–928
- Genin, S., Brito, B., Denny, T. P., and Boucher, C. 2005. Control of the *Ralstonia solanacearum* Type III secretion system (Hrp) genes by the global virulence regulator PhcA. *FEBS Lett.* 579:2077–2081
- Genin, S., and Denny, T. P. 2012. Pathogenomics of the *Ralstonia solanacearum* Species Complex. *Annu. Rev. Phytopathol.* 50:67–89
- Gillette, W. K., and Elkan, G. H. 1996. *Bradyrhizobium* (*Arachis*) sp. strain NC92 contains two *nodD* genes involved in the repression of *nodA* and a *nolA* gene required for the efficient nodulation of host plants. *J. Bacteriol.* 178:2757–66
- Giraud, E., and Fleischman, D. 2004. Nitrogen-fixing symbiosis between photosynthetic bacteria and legumes. *Photosynth. Res.* 82:115–130
- Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.-C., Jaubert, M., Simon, D., Cartieaux, F., Prin, Y., Bena, G., Hannibal, L., Fardoux, J., Kojadinovic, M., Vuillet, L., Lajus, A., Cruveiller, S., Rouy, Z., Mangenot, S., Segurens, B., Dossat, C., Franck, W. L., Chang, W.-S., Saunders, E., Bruce, D., Richardson, P., Normand, P., Dreyfus, B., Pignol, D., Stacey, G., Emerich,

- D., Verméglio, A., Médigue, C., and Sadowsky, M. 2007. Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science*. 316:1307–12
- Gonzalez-Sama, A., Lucas, M. M., de Felipe, M. R., and Pueyo, J. J. 2004. An unusual infection mechanism and nodule morphogenesis in white lupin (*Lupinus albus*). *New Phytol.* 163:371–380
- Goris, J., De Vos, P., Coenye, T., Hoste, B., Janssens, D., Brim, H., Diels, L., Mergeay, M., Kersters, K., and Vandamme, P. 2001. Classification of metal-resistant bacteria from industrial biotopes as *Ralstonia campinensis* sp. nov., *Ralstonia metallidurans* sp. nov. and *Ralstonia basilensis* Steinle et al. 1998 emend. *Int. J. Syst. Evol. Microbiol.* 51:1773–1782
- Griesmann, M., Chang, Y., Liu, X., Song, Y., Haberer, G., Crook, M. B., Billault-Penneteau, B., Lauressergues, D., Keller, J., Imanishi, L., Roswanjaya, Y. P., Kohlen, W., Pujic, P., Battenberg, K., Alloisio, N., Liang, Y., Hilhorst, H., Salgado, M. G., Hocher, V., Gherbi, H., Svistoonoff, S., Doyle, J. J., He, S., Xu, Y., Xu, S., Qu, J., Gao, Q., Fang, X., Fu, Y., Normand, P., Berry, A. M., Wall, L. G., Ané, J.-M., Pawlowski, K., Xu, X., Yang, H., Spannagl, M., Mayer, K. F. X., Wong, G. K.-S., Parniske, M., Delaux, P.-M., and Cheng, S. 2018. Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science* (80-). 361:eaat1743
- Grimault, V., Anais, G., and Prior, P. 1994a. Distribution of *Pseudomonas solanacearum* in the stem tissues of tomato plants with different levels of resistance to bacterial wilt. *Plant Pathol.* 43:663–668
- Grimault, V., Gélie, B., Lemattre, M., Prior, P., and Schmit, J. 1994b. Comparative histology of resistant and susceptible tomato cultivars infected by *Pseudomonas solanacearum*. *Physiol. Mol. Plant Pathol.* 44:105–123
- Guan, S. H., Gris, C., Cruveiller, S., Pouzet, C., Tasse, L., Leru, A., Maillard, A., Médigue, C., Batut, J., Masson-Boivin, C., and Capela, D. 2013. Experimental evolution of nodule intracellular infection in legume symbionts. *ISME J.* 7:1367–77
- Guidot, A., Jiang, W., Ferdy, J.-B., Thébaud, C., Barberis, P., Gouzy, J., and Genin, S. 2014. Multihost Experimental Evolution of the Pathogen *Ralstonia solanacearum* Unveils Genes Involved in Adaptation to Plants. *Mol. Biol. Evol.* 31:2913–2928
- Gyaneshwar, P., Hirsch, A. M., Moulin, L., Chen, W.-M., Elliott, G. N., Bontemps, C., Estrada-de los Santos, P., Gross, E., dos Reis, F. B., Sprent, J. I., Young, J. P. W., and James, E. K. 2011. Legume-Nodulating Betaproteobacteria: Diversity, Host Range, and Future Prospects. *Mol. Plant-Microbe Interact.* 24:1276–1288
- Haag, A. F., Arnold, M. F. F., Myka, K. K., Kerscher, B., Dall'Angelo, S., Zanda, M., Mergaert, P., and Ferguson, G. P. 2013. Molecular insights into bacteroid development during *Rhizobium*–legume symbiosis. *FEMS Microbiol. Rev.* 37:364–383
- Haag, A. F., Wehmeier, S., Beck, S., Marlow, V. L., Fletcher, V., James, E. K., and Ferguson, G. P. 2009. The Sinorhizobium meliloti LpxXL and AcpXL proteins play important roles in bacteroid development within alfalfa. *J. Bacteriol.* 191:4681–6
- Hayward, A. C. 1991. Biology and epidemiology of bacterial wilt caused by *pseudomonas solanacearum*. *Annu. Rev. Phytopathol.* 29:65–87
- He, X., and Zhang, J. 2005. Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics.* 169:1157–64
- Heath, K. D., and Tiffin, P. 2009. STABILIZING MECHANISMS IN A LEGUME-RHIZOBIUM MUTUALISM. *Evolution* (N. Y). 63:652–662
- Hirsch, A. M., Lum, M. R., and Downie, J. A. 2001. What makes the rhizobia-legume symbiosis so special? *Plant Physiol.* 127:1484–92
- Hirsch, A. M., and Smith, C. A. 1987. Effects of *Rhizobium meliloti* nif and fix mutants on alfalfa root nodule development. *J. Bacteriol.* 169:1137–1146
- Huisman, R., Ovchinnikova, E., Bisseling, T., and Limpens, E. 2012. Endocytic Accommodation of Microbes in Plants. Pages 271–295 in: *Endocytosis in Plants*, Springer Berlin Heidelberg, Berlin, Heidelberg.
- John, M., Schmidt, J., Wieneke, U., Krüssmann, H.-D., and Schell, J. 1988. Transmembrane orientation

- and receptor-like structure of the *Rhizobium meliloti* common nodulation protein NodC. *EMBO J.* 7:583–588
- Johnston, A. W. B., Beynon, J. L., Buchanan-Wollaston, A. V., Setchell, S. M., Hirsch, P. R., and Beringer, J. E. 1978. High frequency transfer of nodulating ability between strains and species of *Rhizobium*. *Nature.* 276:634–636
- Jones, E. I., Afkhami, M. E., Akçay, E., Bronstein, J. L., Bshary, R., Frederickson, M. E., Heath, K. D., Hoeksema, J. D., Ness, J. H., Pankey, M. S., Porter, S. S., Sachs, J. L., Scharnagl, K., and Friesen, M. L. 2015. Cheaters must prosper: reconciling theoretical and empirical perspectives on cheating in mutualism. N. van Dam, ed. *Ecol. Lett.* 18:1270–1284
- Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., and Walker, G. C. 2007. How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. *Nat. Rev. Microbiol.* 5:619–33
- Kaló, P., Seres, A., Taylor, S. A., Jakab, J., Kevei, Z., Kereszt, A., Endre, G., Ellis, T. H. N., and Kiss, G. B. 2004. Comparative mapping between *Medicago sativa* and *Pisum sativum*. *Mol. Genet. Genomics.* 272:235–246
- Kao, C. C., Barlow, E., and Sequeira, L. 1992. Extracellular polysaccharide is required for wild-type virulence of *Pseudomonas solanacearum*. *J. Bacteriol.* 174:1068–71
- Karafin, M., Romagnoli, M., Fink, D. L., Howard, T., Rau, R., Milstone, A. M., and Carroll, K. C. 2010. Fatal infection caused by *Cupriavidus gilardii* in a child with aplastic anemia. *J. Clin. Microbiol.* 48:1005–7
- Kelly, S. A., Czech, P. P., Wight, J. T., Blank, K. M., and Garland, T. 2006. Experimental evolution and phenotypic plasticity of hindlimb bones in high-activity house mice. *J. Morphol.* 267:360–374
- Kereszt, A., Mergaert, P., and Kondorosi, E. 2011. Bacteroid Development in Legume Nodules: Evolution of Mutual Benefit or of Sacrificial Victims? *Mol. Plant-Microbe Interact.* 24:1300–1309
- Kiers, E. T., Rousseau, * Robert A., and Denison, R. F. 2006. *Evolutionary ecology research.* Evolutionary Ecology.
- Kiers, E. T., Rousseau, R. A., West, S. A., and Denison, R. F. 2003. Host sanctions and the legume-rhizobium mutualism. *Nature.* 425:78–81
- Klonowska, A., Chaintreuil, C., Tisseyre, P., Miché, L., Melkonian, R., Ducouso, M., Laguerre, G., Brunel, B., and Moulin, L. 2012. Biodiversity of *Mimosa pudica* rhizobial symbionts (*Cupriavidus taiwanensis*, *Rhizobium mesoamericanum*) in New Caledonia and their adaptation to heavy metal-rich soils. *FEMS Microbiol. Ecol.* 81:618–635
- Klonowska, A., Melkonian, R., Miché, L., Tisseyre, P., and Moulin, L. 2018. Transcriptomic profiling of *Burkholderia phymatum* STM815, *Cupriavidus taiwanensis* LMG19424 and *Rhizobium mesoamericanum* STM3625 in response to *Mimosa pudica* root exudates illuminates the molecular basis of their nodulation competitiveness and symbiotic ev. *BMC Genomics.* 19:105
- Kneip, C., Lockhart, P., Voß, C., and Maier, U.-G. 2007. Nitrogen fixation in eukaryotes – New models for symbiosis. *BMC Evol. Biol.* 7:55
- Kovach, M. E., Elzer, P. H., Steven Hill, D., Robertson, G. T., Farris, M. A., Roop, R. M., and Peterson, K. M. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene.* 166:175–176
- Krishnakumar, V., Kim, M., Rosen, B. D., Karamycheva, S., Bidwell, S. L., Tang, H., and Town, C. D. 2015. MTGD: The *Medicago truncatula* Genome Database. *Plant Cell Physiol.* 56:e1–e1
- Krishnan, H. B., Lorio, J., Kim, W. S., Jiang, G., Kim, K. Y., DeBoer, M., and Pueppke, S. G. 2003. Extracellular Proteins Involved in Soybean Cultivar-Specific Nodulation Are Associated with Pilus-Like Surface Appendages and Exported by a Type III Protein Secretion System in *Sinorhizobium fredii* USDA257. *Mol. Plant-Microbe Interact.* 16:617–625
- Krishnan, H. B., and Pueppke, S. G. 1991. Sequence and analysis of the nodABC region of *Rhizobium fredii* USDA257, a nitrogen-fixing symbiont of soybean and other legumes. *Mol. Plant. Microbe Interact.* 4:512–520
- Ktari, A., Nouioui, I., Furnholm, T., Swanson, E., Ghodhbane-Gtari, F., Tisa, L. S., and Gtari, M. 2017. Permanent draft genome sequence of *Frankia* sp. NRRL B-16219 reveals the presence of canonical nod genes, which are highly homologous to those detected in Candidatus *Frankia* Dg1

- genome. *Stand. Genomic Sci.* 12:51
- Kumar, K., Mella-Herrera, R. A., and Golden, J. W. 2010. Cyanobacterial heterocysts. *Cold Spring Harb. Perspect. Biol.* 2:a000315
- Laguerre, G., Heulin-Gotty, K., Brunel, B., Klonowska, A., Le Quéré, A., Tillard, P., Prin, Y., Cleyet-Marel, J.-C., and Lepetit, M. 2012. Local and systemic N signaling are involved in *Medicago truncatula* preference for the most efficient *Sinorhizobium* symbiotic partners. *New Phytol.* 195:437–49
- De Lajudie, P., Willems, A., Pot, B., Dewettinck, D., Maestrojuan, G., Neyra, M., Collins, M. D., Dreyfus, B., Kersters, K., and Gillis, M. 1994. Polyphasic Taxonomy of Rhizobia: Emendation of the Genus *Sinorhizobium* and Description of *Sinorhizobium meliloti* comb. nov., *Sinorhizobium saheli* sp. nov., and *Sinorhizobium teranga* sp. nov. *Int. J. Syst. Bacteriol.* 44:715–733
- Lal, S., and Tabacchioni, S. 2009. Ecology and biotechnological potential of *Paenibacillus polymyxa*: a minireview. *Indian J. Microbiol.* 49:2–10
- Latysheva, N., Junker, V. L., Palmer, W. J., Codd, G. A., and Barker, D. 2012. The evolution of nitrogen fixation in cyanobacteria. *Bioinformatics.* 28:603–606
- Lee, C. C., Hu, Y., and Ribbe, M. W. 2010. Vanadium nitrogenase reduces CO. *Science.* 329:642
- Lenski, R. E., Rose, M. R., Simpson, S. C., and Tadler, S. C. 1991. Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. *Am. Nat.* 138:1315–1341
- Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.-P., Ané, J.-M., Lauber, E., Bisseling, T., Dénarié, J., Rosenberg, C., and Debelle, F. 2004. A putative Ca²⁺ and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science.* 303:1361–4
- Lewin, A., Rosenberg, C., H. Meyer, z. A., Wong, C. H., Nelson, L., Manen, J.-F., Stanley, J., Dowling, D. N., Denarie, J., and Broughton, W. J. 1987. Multiple host-specificity loci of the broad host-range *Rhizobium* sp. NGR234 selected using the widely compatible legume *Vigna unguiculata*. *Plant Mol. Biol.* 8:447–459
- Lindström, K. 1989. *Rhizobium galegae*, a New Species of Legume Root Nodule Bacteria. *Int. J. Syst. Bacteriol.* 39:365–367
- Ling, J., Zheng, H., Katzianer, D. S., Wang, H., Zhong, Z., and Zhu, J. 2013. Applying Reversible Mutations of Nodulation and Nitrogen-Fixation Genes to Study Social Cheating in *Rhizobium etli*-Legume Interaction D. Ballhorn, ed. *PLoS One.* 8:e70138
- Liu, C.-W., and Murray, J. 2016. The Role of Flavonoids in Nodulation Host-Range Specificity: An Update. *Plants.* 5:33
- Liu, X. Y., Wu, W., Wang, E. T., Zhang, B., Macdermott, J., and Chen, W. X. 2011. Phylogenetic relationships and diversity of α -rhizobia associated with *Mimosa* species grown in Sishuangbanna, China. *Int. J. Syst. Evol. Microbiol.* 61:334–342
- Loh, J., Garcia, M., and Stacey, G. 1997. NodV and NodW, a second flavonoid recognition system regulating nod gene expression in *Bradyrhizobium japonicum*. *J. Bacteriol.* 179:3013–20
- Loh, J., Lohar, D. P., Andersen, B., and Stacey, G. 2002. A two-component regulator mediates population-density-dependent expression of the *Bradyrhizobium japonicum* nodulation genes. *J. Bacteriol.* 184:1759–66
- Madsen, E. B., Antolín-Llovera, M., Grossmann, C., Ye, J., Vieweg, S., Broghammer, A., Krusell, L., Radutoiu, S., Jensen, O. N., Stougaard, J., and Parniske, M. 2011. Autophosphorylation is essential for the in vivo function of the *Lotus japonicus* Nod factor receptor 1 and receptor-mediated signalling in cooperation with Nod factor receptor 5. *Plant J.* 65:404–417
- Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N., and Stougaard, J. 2003. A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature.* 425:637–640
- Madsen, L. H., Tirichine, L., Jurkiewicz, A., Sullivan, J. T., Heckmann, A. B., Bek, A. S., Ronson, C. W., James, E. K., and Stougaard, J. 2010. The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat. Commun.* 1:1–12
- Magallon, S., Crane, P. R., and Herendeen, P. S. 1999. Phylogenetic Pattern, Diversity, and

- Diversification of Eudicots. *Ann. Missouri Bot. Gard.* 86:297
- Maillet, F., Debellé, F., and Dénarié, J. 1990. Role of the *nodD* and *syrM* genes in the activation of the regulatory gene *nodD3*, and of the common and host-specific *nod* genes of *Rhizobium meliloti*. *Mol. Microbiol.* 4:1975–1984
- Mano, H., Fujii, T., Sumikawa, N., Hiwatashi, Y., and Hasebe, M. 2014. Development of an *Agrobacterium*-Mediated Stable Transformation Method for the Sensitive Plant *Mimosa pudica*. B.A. Vinatzer, ed. *PLoS One.* 9:e88611
- Marchetti, M., Capela, D., Glew, M., Cruveiller, S., Chane-Woon-Ming, B., Gris, C., Timmers, T., Poinot, V., Gilbert, L. B., Heeb, P., Médigue, C., Batut, J., and Masson-Boivin, C. 2010. Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biol.* 8:e1000280
- Marchetti, M., Catrice, O., Batut, J., and Masson-Boivin, C. 2011. *Cupriavidus taiwanensis* bacteroids in *Mimosa pudica* Indeterminate nodules are not terminally differentiated. *Appl. Environ. Microbiol.* 77:2161–4
- Marchetti, M., Clerissi, C., Yousfi, Y., Gris, C., Bouchez, O., Rocha, E., Cruveiller, S., Jauneau, A., Capela, D., and Masson-Boivin, C. 2016. Experimental evolution of rhizobia may lead to either extra- or intracellular symbiotic adaptation depending on the selection regime. *Mol. Ecol.*
- Marchetti, M., Jauneau, A., Capela, D., Remigi, P., Gris, C., Batut, J., and Masson-Boivin, C. 2014. Shaping Bacterial Symbiosis With Legumes by Experimental Evolution. *Mol. Plant-Microbe Interact.* 27:956–964
- Marco, D. E., Pérez-Arnedo, R., Hidalgo-Perea, Á., Olivares, J., Ruiz-Sainz, J. E., and Sanjuán, J. 2009. A mechanistic molecular test of the plant-sanction hypothesis in legume–rhizobia mutualism. *Acta Oecologica.* 35:664–667
- Martínez-Romero, E. 1994. Recent developments in *Rhizobium* taxonomy. Pages 11–20 in: *Symbiotic Nitrogen Fixation*, Springer Netherlands, Dordrecht.
- Martínez-Romero, J. C., Ormeño-Orrillo, E., Rogel, M. A., López-López, A., and Martínez-Romero, E. 2010. Trends in Rhizobial Evolution and Some Taxonomic Remarks. Pages 301–315 in: *Evolutionary Biology – Concepts, Molecular and Morphological Evolution*, Springer Berlin Heidelberg, Berlin, Heidelberg.
- Marx, C. J., and Lidstrom, M. E. 2002. Broad-Host-Range *cre-lox* System for Antibiotic Marker Recycling in Gram-Negative Bacteria. *Biotechniques.* 33:1062–1067
- Masson-Boivin, C., Giraud, E., Perret, X., and Batut, J. 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol.* 17:458–66
- Masson-Boivin, C., and Sachs, J. L. 2018. Symbiotic nitrogen fixation by rhizobia — the roots of a success story. *Curr. Opin. Plant Biol.* 44:7–15
- Melkonian, R., Moulin, L., Béna, G., Tisseyre, P., Chaintreuil, C., Heulin, K., Rezkallah, N., Klonowska, A., Gonzalez, S., Simon, M., Chen, W.-M., James, E. K., and Laguerre, G. 2014. The geographical patterns of symbiont diversity in the invasive legume *Mimosa pudica* can be explained by the competitiveness of its symbionts and by the host genotype. *Environ. Microbiol.* 16:2099–2111
- Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., and Kondorosi, E. 2003. A Novel Family in *Medicago truncatula* Consisting of More Than 300 Nodule-Specific Genes Coding for Small, Secreted Polypeptides with Conserved Cysteine Motifs. *PLANT Physiol.* 132:161–173
- Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.-E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., and Kondorosi, E. 2006. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* 103:5230–5
- Messinese, E., Mun, J.-H., Yeun, L. H., Jayaraman, D., Rougé, P., Barre, A., Loughon, G., Schornack, S., Bono, J.-J., Cook, D. R., and Ané, J.-M. 2007. A Novel Nuclear Protein Interacts With the Symbiotic DMI3 Calcium- and Calmodulin-Dependent Protein Kinase of *Medicago truncatula*. *Mol. Plant-Microbe Interact.* 20:912–921
- Minamisawa, K., Nishioka, K., Miyaki, T., Ye, B., Miyamoto, T., You, M., Saito, A., Saito, M., Barraquio, W. L., Teaumroong, N., Sein, T., and Sato, T. 2004. Anaerobic Nitrogen-Fixing Consortia

- Consisting of Clostridia Isolated from Gramineous Plants. *Appl. Environ. Microbiol.* 70:3096–3102
- Mishra, R. P. N., Tisseyre, P., Melkonian, R., Chaintreuil, C., Miché, L., Klonowska, A., Gonzalez, S., Bena, G., Laguerre, G., and Moulin, L. 2012. Genetic diversity of *Mimosa pudica* rhizobial symbionts in soils of French Guiana: investigating the origin and diversity of *Burkholderia phymatum* and other beta-rhizobia. *FEMS Microbiol. Ecol.* 79:487–503
- Miwa, H., Sun, J., Oldroyd, G. E. D., and Downie, J. A. 2006. Analysis of Nod-Factor-Induced Calcium Signaling in Root Hairs of Symbiotically Defective Mutants of *Lotus japonicus*. *Mol. Plant-Microbe Interact.* 19:914–923
- Moulin, L., Munive, A., Dreyfus, B., and Boivin-Masson, C. 2001. Nodulation of legumes by members of the β -subclass of Proteobacteria. *Nature.* 411:948–950
- Nakagawa, T., Kaku, H., Shimoda, Y., Sugiyama, A., Shimamura, M., Takanashi, K., Yazaki, K., Aoki, T., Shibuya, N., and Kouchi, H. 2011. From defense to symbiosis: limited alterations in the kinase domain of LysM receptor-like kinases are crucial for evolution of legume-Rhizobium symbiosis. *Plant J.* 65:169–180
- Nandasena, K. G., O'hara, G. W., Tiwari, R. P., and Howieson, J. G. 2006. Rapid in situ evolution of nodulating strains for *Biserrula pelecinus* L. through lateral transfer of a symbiosis island from the original mesorhizobial inoculant. *Appl. Environ. Microbiol.* 72:7365–7
- Nandasena, K. G., O'Hara, G. W., Tiwari, R. P., Sezmiş, E., and Howieson, J. G. 2007. In situ lateral transfer of symbiosis islands results in rapid evolution of diverse competitive strains of mesorhizobia suboptimal in symbiotic nitrogen fixation on the pasture legume *Biserrula pelecinus* L. *Environ. Microbiol.* 9:2496–2511
- Ndoye, I., de Billy, F., Vasse, J., Dreyfus, B., and Truchet, G. 1994. Root nodulation of *Sesbania rostrata*. *J. Bacteriol.* 176:1060–8
- Newton, W. E. 2000. Nitrogen Fixation in Perspective. Pages 3–8 in: *Nitrogen Fixation: From Molecules to Crop Productivity*, Kluwer Academic Publishers, Dordrecht.
- Nguyen, T. Van, Wibberg, D., Battenberg, K., Blom, J., Vanden Heuvel, B., Berry, A. M., Kalinowski, J., and Pawlowski, K. 2016. An assemblage of *Frankia* Cluster II strains from California contains the canonical nod genes and also the sulfotransferase gene nodH. *BMC Genomics.* 17:796
- Nick, G., Jussila, M., Hoste, B., Niemi, R. M., Kaijalainen, S., de Lajudie, P., Gillis, M., de Bruijn, F. J., and Lindström, K. 1999. Rhizobia Isolated from Root Nodules of Tropical Leguminous Trees Characterized Using DNA-DNA dot-blot Hybridisation and rep-PCR Genomic Fingerprinting. *Syst. Appl. Microbiol.* 22:287–299
- Okazaki, S., Kaneko, T., Sato, S., and 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., Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., Gully, D., Arrighi, J.-F., Furuta, N., Miwa, H., Yasuda, M., Nouwen, N., Teumroong, N., and Giraud, E. 2016. Rhizobium–legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. *ISME J.* 10:64–74
- Oldroyd, G. E. D. 2013. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* 11:252–63
- Oldroyd, G. E. D., and Downie, J. A. 2008. Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annu. Rev. Plant Biol.* 59:519–46
- Oldroyd, G. E. D., and Long, S. R. 2003. Identification and characterization of nodulation-signaling pathway 2, a gene of *Medicago truncatula* involved in Nod actor signaling. *Plant Physiol.* 131:1027–32
- Oono, R., Anderson, C. G., and Denison, R. F. 2011. Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. *Proc. Biol. Sci.* 278:2698–703
- Ormeño-Orrillo, E., Servín-Garcidueñas, L. E., Imperial, J., Rey, L., Ruiz-Argueso, T., and Martínez-Romero, E. 2013. Phylogenetic evidence of the transfer of nodZ and nolL genes from *Bradyrhizobium* to other rhizobia. *Mol. Phylogenet. Evol.* 67:626–630

- Orr, C. H., James, A., Leifert, C., Cooper, J. M., and Cummings, S. P. 2011. Diversity and activity of free-living nitrogen-fixing bacteria and total bacteria in organic and conventionally managed soils. *Appl. Environ. Microbiol.* 77:911–9
- Patriarca, E. J., Tatè, R., and Iaccarino, M. 2002. Key role of bacterial NH₄(+) metabolism in Rhizobium-plant symbiosis. *Microbiol. Mol. Biol. Rev.* 66:203–22
- Pawlowski, K., and Demchenko, K. N. 2012. The diversity of actinorhizal symbiosis. *Protoplasma.* 249:967–979
- Perret, X., Freiberg, C., Rosenthal, A., Broughton, W. J., and Fellay, R. 1999. High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234. *Mol. Microbiol.* 32:415–425
- Perret, X., Staehelin, C., and Broughton, W. J. 2000. Molecular basis of symbiotic promiscuity. *Microbiol. Mol. Biol. Rev.* 64:180–201
- Perrier, A., Peyraud, R., Rengel, D., Barlet, X., Lucasson, E., Gouzy, J., Peeters, N., Genin, S., and Guidot, A. 2016. Enhanced in planta Fitness through Adaptive Mutations in EfpR, a Dual Regulator of Virulence and Metabolic Functions in the Plant Pathogen *Ralstonia solanacearum* D. Desveaux, ed. *PLoS Pathog.* 12:e1006044
- Perry, J. A., Wang, T. L., Welham, T. J., Gardner, S., Pike, J. M., Yoshida, S., and Parniske, M. 2003. A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiol.* 131:866–71
- Persson, T., Battenberg, K., Demina, I. V., Vigil-Stenman, T., Vanden Heuvel, B., Pujic, P., Facciotti, M. T., Wilbanks, E. G., O'Brien, A., Fournier, P., Cruz Hernandez, M. A., Mendoza Herrera, A., Médigue, C., Normand, P., Pawlowski, K., and Berry, A. M. 2015. Candidatus *Frankia Datiscae* Dg1, the Actinobacterial Microsymbiont of *Datisca glomerata*, Expresses the Canonical nod Genes nodABC in Symbiosis with Its Host Plant *F. Börnke*, ed. *PLoS One.* 10:e0127630
- Peyraud, R., Cottret, L., Marmiesse, L., Gouzy, J., and Genin, S. 2016. A Resource Allocation Trade-Off between Virulence and Proliferation Drives Metabolic Versatility in the Plant Pathogen *Ralstonia solanacearum* D. Desveaux, ed. *PLoS Pathog.* 12:e1005939
- Poinsot, V., Crook, M. B., Erdn, S., Maillet, F., Bascaules, A., and Ané, J.-M. 2016. New insights into Nod factor biosynthesis: Analyses of chitooligomers and lipo-chitooligomers of *Rhizobium* sp. IRBG74 mutants. *Carbohydr. Res.* 434:83–93
- Poueymiro, M., and Genin, S. 2009. Secreted proteins from *Ralstonia solanacearum*: a hundred tricks to kill a plant. *Curr. Opin. Microbiol.* 12:44–52
- Prell, J., and Poole, P. 2006. Metabolic changes of rhizobia in legume nodules. *Trends Microbiol.* 14:161–168
- Price, P. A., Tanner, H. R., Dillon, B. A., Shabab, M., Walker, G. C., and Griffiths, J. S. 2015. Rhizobial peptidase HrrP cleaves host-encoded signaling peptides and mediates symbiotic compatibility. *Proc. Natl. Acad. Sci. U. S. A.* 112:15244–9
- Pueppke, S. G., and Broughton, W. J. 1999. *Rhizobium* sp. Strain NGR234 and *R. fredii* USDA257 Share Exceptionally Broad, Nested Host Ranges. *Mol. Plant-Microbe Interact.* 12:293–318
- Rai, A. N., Söderbäck, E., and Bergman, B. 2000. Cyanobacterium–plant symbioses. *New Phytol.* 147:449–481
- Rastogi, S., and Liberles, D. A. 2005. Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol. Biol.* 5:28
- Rees, D. C., and Howard, J. B. 2000. Nitrogenase: standing at the crossroads. *Curr. Opin. Chem. Biol.* 4:559–566
- Regus, J. U., Quides, K. W., O'Neill, M. R., Suzuki, R., Savory, E. A., Chang, J. H., and Sachs, J. L. 2017. Cell autonomous sanctions in legumes target ineffective rhizobia in nodules with mixed infections. *Am. J. Bot.* 104:1299–1312
- Rehder, D. 2000. Vanadium nitrogenase. *J. Inorg. Biochem.* 80:133–136
- Reid, D. E., Ferguson, B. J., Hayashi, S., Lin, Y.-H., and Gresshoff, P. M. 2011. Molecular mechanisms controlling legume autoregulation of nodulation. *Ann. Bot.* 108:789–795
- Remigi, P., Capela, D., Clerissi, C., Tasse, L., Torchet, R., Bouchez, O., Batut, J., Cruveiller, S., Rocha, E.

- P. C., and Masson-Boivin, C. 2014. Transient hypermutagenesis accelerates the evolution of legume endosymbionts following horizontal gene transfer. *J.L. Dangl, ed. PLoS Biol.* 12:e1001942
- Remigi, P., Masson-Boivin, C., and Rocha, E. P. C. 2019. Experimental Evolution as a Tool to Investigate Natural Processes and Molecular Functions. *Trends Microbiol.*
- Remigi, P., Zhu, J., Young, J. P. W., and Masson-Boivin, C. 2016. Symbiosis within Symbiosis: Evolving Nitrogen-Fixing Legume Symbionts. *Trends Microbiol.* 24:63–75
- van Rhijn, P. J., Feys, B., Verreth, C., and Vanderleyden, J. 1993. Multiple copies of nodD in *Rhizobium tropici* CIAT899 and BR816. *J. Bacteriol.* 175:438–47
- Riely, B. K., Lougnon, G., Ané, J.-M., and Cook, D. R. 2007. The symbiotic ion channel homolog DMI1 is localized in the nuclear membrane of *Medicago truncatula* roots. *Plant J.* 49:208–216
- Rivas, R., Velázquez, E., Willems, A., Vizcaíno, N., Subba-Rao, N. S., Mateos, P. F., Gillis, M., Dazzo, F. B., and Martínez-Molina, E. 2002. A new species of *Devosia* that forms a unique nitrogen-fixing root-nodule symbiosis with the aquatic legume *Neptunia natans* (L.f.) druce. *Appl. Environ. Microbiol.* 68:5217–22
- Rogel, M. A., Ormeño-Orrillo, E., and Martínez Romero, E. 2011. Symbiovars in rhizobia reflect bacterial adaptation to legumes. *Syst. Appl. Microbiol.* 34:96–104
- Saad, M. M., Crèvecoeur, M., Masson-Boivin, C., and Perret, X. 2012. The type 3 protein secretion system of *Cupriavidus taiwanensis* strain LMG19424 compromises symbiosis with *Leucaena leucocephala*. *Appl. Environ. Microbiol.* 78:7476–9
- Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., Billault, A., Brottier, P., Camus, J. C., Cattolico, L., Chandler, M., Choisne, N., Claudel-Renard, C., Cunnac, S., Demange, N., Gaspin, C., Lavie, M., Moisan, A., Robert, C., Saurin, W., Schiex, T., Siguier, P., Thébault, P., Whalen, M., Wincker, P., Levy, M., Weissenbach, J., and Boucher, C. A. 2002. Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature.* 415:497. 6p. 3 Diagrams
- Sanjuan, J., Grob, P., Gottfert, M., Hennecke, H., and Stacey, G. 1994. NodW Is Essential for Full Expression of the Common Nodulation Genes in *Bradyrhizobium japonicum*. *Mol. Plant-Microbe Interact.* 7:364
- Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., Sasamoto, S., Watanabe, A., Ono, A., Kawashima, K., Fujishiro, T., Katoh, M., Kohara, M., Kishida, Y., Minami, C., Nakayama, S., Nakazaki, N., Shimizu, Y., Shinpo, S., Takahashi, C., Wada, T., Yamada, M., Ohmido, N., Hayashi, M., Fukui, K., Baba, T., Nakamichi, T., Mori, H., and Tabata, S. 2008. Genome Structure of the Legume, *Lotus japonicus*. *DNA Res.* 15:227–239
- Schlaman, H. R., Spaink, H. P., Okker, R. J., and Lugtenberg, B. J. 1989. Subcellular localization of the nodD gene product in *Rhizobium leguminosarum*. *J. Bacteriol.* 171:4686–93
- Schlötterer, C., Kofler, R., Versace, E., Tobler, R., and Franssen, S. U. 2015. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity (Edinb).* 114:431–440
- Schneider, K., Gollan, U., Drottboom, M., Selsemeier-Voigt, S., and Müller, A. 1997. Comparative Biochemical Characterization of the Iron-Only Nitrogenase and the Molybdenum Nitrogenase from *Rhodobacter capsulatus*. *Eur. J. Biochem.* 244:789–800
- Schultze, M., Quiclet-Sire, B., Kondorosi, E., Virelizer, H., Glushka, J. N., Endre, G., Géro, S. D., and Kondorosi, A. 1991. *Rhizobium meliloti* produces a family of sulfated lipooligosaccharides exhibiting different degrees of plant host specificity. *Proc. Natl. Acad. Sci. U. S. A.* 89:192–6
- Seefeldt, L. C., Hoffman, B. M., and Dean, D. R. 2009. Mechanism of Mo-Dependent Nitrogenase. *Annu. Rev. Biochem.* 78:701–722
- Sellstedt, A., and Richau, K. H. 2013. Aspects of nitrogen-fixing *Actinobacteria*, in particular free-living and symbiotic *Frankia*. *FEMS Microbiol. Lett.* 342:179–186
- Simms, E. L., Taylor, D. L., Povich, J., Shefferson, R. P., Sachs, J. ., Urbina, M., and Tausczik, Y. 2006. An empirical test of partner choice mechanisms in a wild legume–rhizobium interaction. *Proc. R. Soc. B Biol. Sci.* 273:77–81
- Simon, M. F., Grether, R., de Queiroz, L. P., Särkinen, T. E., Dutra, V. F., and Hughes, C. E. 2011a. The

- evolutionary history of *Mimosa* (Leguminosae): Toward a phylogeny of the sensitive plants. *Am. J. Bot.* 98:1201–1221
- Simon, M. F., Grether, R., de Queiroz, L. P., Särkinen, T. E., Dutra, V. F., and Hughes, C. E. 2011b. The evolutionary history of *Mimosa* (Leguminosae): Toward a phylogeny of the sensitive plants. *Am. J. Bot.* 98:1201–1221
- Singleton, P. W., and Stockinger, K. R. 1983. Compensation against Ineffective Nodulation in Soybean. *Crop Sci.* 23:69
- Skorupska, A., Janczarek, M., Marczak, M., Mazur, A., and Król, J. 2006. Rhizobial exopolysaccharides: Genetic control and symbiotic functions. *Microb. Cell Fact.* 5:7
- Snoeck, C., Luyten, E., Poinsoot, V., Savagnac, A., Vanderleyden, J., and Promé, J.-C. 2001. Rhizobium sp. BR816 Produces a Complex Mixture of Known and Novel Lipochitooligosaccharide Molecules. *Mol. Plant-Microbe Interact.* 14:678–684
- Soltis, D. E., Soltis, P. S., Morgan, D. R., Swensen, S. M., Mullin, B. C., Dowd, J. M., and Martin, P. G. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc. Natl. Acad. Sci. U. S. A.* 92:2647–51
- Songwattana, P., Noisangiam, R., Teamtisong, K., Prakamhang, J., Teulet, A., Tittabutr, P., Piromyou, P., Boonkerd, N., Giraud, E., and Teaumroong, N. 2017. Type 3 Secretion System (T3SS) of *Bradyrhizobium* sp. DOA9 and Its Roles in Legume Symbiosis and Rice Endophytic Association. *Front. Microbiol.* 8:1810
- Soria-Díaz, M. E., Rodrí-guez-Carvajal, M. A., Tejero-Mateo, P., Espartero, J. L., Morón, B., Sousa, C., Megí-as, M., Thomas-Oates, J., and Gil-Serrano, A. M. 2006. Structural determination of the Nod factors produced by *Rhizobium gallicum* bv. *gallicum* R602. *FEMS Microbiol. Lett.* 255:164–173
- Spaink, H. P. 2000. Root Nodulation and Infection Factors Produced by Rhizobial Bacteria. *Annu. Rev. Microbiol.* 54:257–288
- Sprent, J. I. 2009. *Legume nodulation : a global perspective*. Wiley-Blackwell.
- Subramanian, S., Stacey, G., and Yu, O. 2007. Distinct, crucial roles of flavonoids during legume nodulation. *Trends Plant Sci.* 12:282–285
- Sullivan, J. T., Patrick, H. N., Lowther, W. L., Scott, D. B., and Ronson, C. W. 1995a. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. *Proc. Natl. Acad. Sci. U. S. A.* 92:8985–9
- Sullivan, J. T., Patrick, H. N., Lowther, W. L., Scott, D. B., and Ronson, C. W. 1995b. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. *Proc. Natl. Acad. Sci. U. S. A.* 92:8985–8989
- Sullivan, J. T., and Ronson, C. W. 1998. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a *phe-tRNA* gene. *Proc. Natl. Acad. Sci.* 95:5145–5149
- Swensen, S. M. 1996. The evolution of actinorhizal symbioses: Evidence for multiple origins of the symbiotic association. *Am. J. Bot.* 83:1503–1512
- Sy, A., Giraud, E., Jourand, P., Garcia, N., Willems, A., de Lajudie, P., Prin, Y., Neyra, M., Gillis, M., Boivin-Masson, C., and Dreyfus, B. 2001. Methylophilic *Methylobacterium* bacteria nodulate and fix nitrogen in symbiosis with legumes. *J. Bacteriol.* 183:214–20
- Takeda, N., Maekawa, T., and Hayashi, M. 2012. Nuclear-localized and deregulated calcium- and calmodulin-dependent protein kinase activates rhizobial and mycorrhizal responses in *Lotus japonicus*. *Plant Cell.* 24:810–22
- Tang, M., and Capela, D. 2020. Rhizobium diversity in the light of evolution. *Adv. Bot. Res.* 94:251–288
- Tans-Kersten, J., Brown, D., and Allen, C. 2004. Swimming Motility, a Virulence Trait of *Ralstonia solanacearum*, Is Regulated by FlhDC and the Plant Host Environment. *Mol. Plant-Microbe Interact.* 17:686–695
- Taylor, E. J., Gloster, T. M., Turkenburg, J. P., Vincent, F., Brzozowski, A. M., Dupont, C., Shareck, F., Centeno, M. S. J., Prates, J. A. M., Puchart, V., Ferreira, L. M. A., Fontes, C. M. G. A., Biely, P., and Davies, G. J. 2006. Structure and activity of two metal ion-dependent acetylxyylan esterases

- involved in plant cell wall degradation reveals a close similarity to peptidoglycan deacetylases. *J. Biol. Chem.* 281:10968–75
- Teotónio, H., Estes, S., Phillips, P. C., and Baer, C. F. 2017. Experimental Evolution with *Caenorhabditis* Nematodes. *Genetics.* 206:691–716
- Terefework, Z., Nick, G., Suomalainen, S., Paulin, L., and Lindstrom, K. 1998. Phylogeny of *Rhizobium galegae* with respect to other rhizobia and agrobacteria. *Int. J. Syst. Bacteriol.* 48:349–356
- Tilden, E. B., and Hudson, C. S. 1942. Preparation and Properties of the Amylases Produced by *Bacillus macerans* and *Bacillus polymyxa*. *J. Bacteriol.* 43:527–44
- Timmers, A. C., Auriac, M. C., de Billy, F., and Truchet, G. 1998. Nod factor internalization and microtubular cytoskeleton changes occur concomitantly during nodule differentiation in alfalfa. *Development.* 125:339–49
- Turner, S. L., and Young, J. P. W. 2000. The Glutamine Synthetases of Rhizobia: Phylogenetics and Evolutionary Implications. *Mol. Biol. Evol.* 17:309–319
- Valls, M., Genin, S., and Boucher, C. 2006. Integrated Regulation of the Type III Secretion System and Other Virulence Determinants in *Ralstonia solanacearum*. *PLoS Pathog.* 2:e82
- Vandamme, P., and Coenye, T. 2004. Taxonomy of the genus *Cupriavidus*: a tale of lost and found. *Int. J. Syst. Evol. Microbiol.* 54:2285–2289
- Vasse, J. 1995. Microscopic Studies of Intercellular Infection and Protoxylem Invasion of Tomato Roots by *Pseudomonas solanacearum*. *Mol. Plant-Microbe Interact.* 8:241
- Vasse, J., de Billy, F., Camut, S., and Truchet, G. 1990. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J. Bacteriol.* 172:4295–306
- van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D., Roswanjaya, Y., Wardhani, T. A. K., Kalhor, M. S., Jansen, J., van den Hoogen, J., Güngör, B., Hartog, M., Hontelez, J., Verver, J., Yang, W.-C., Schijlen, E., Repin, R., Schilthuisen, M., Schranz, M. E., Heidstra, R., Miyata, K., Fedorova, E., Kohlen, W., Bisseling, T., Smit, S., and Geurts, R. 2018. Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. *Proc. Natl. Acad. Sci. U. S. A.* 115:E4700–E4709
- Verma, S. C., Chowdhury, S. P., and Tripathi, A. K. 2004. Phylogeny based on 16S rDNA and *nifH* sequences of *Ralstonia taiwanensis* strains isolated from nitrogen-fixing nodules of *Mimosa pudica*, in India. *Can. J. Microbiol.* 50:313–322
- Ververidis, F., Trantas, E., Douglas, C., Vollmer, G., Kretzschmar, G., and Panopoulos, N. 2007. Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health. *Biotechnol. J.* 2:1214–1234
- Vlassak, K. M., Luyten, E., Verreth, C., Van Rhijn, P., Bisseling, T., and Vanderleyden, J. 1998. The *Rhizobium* sp. BR816 nodO gene can function as a determinant for nodulation of *Leucaena leucocephala*, *Phaseolus vulgaris*, and *Trifolium repens* by a diversity of *Rhizobium* spp. *Mol. Plant-Microbe Interact.* 11:383–392
- Walker, S. A., and Downie, J. A. 2000. Entry of *Rhizobium leguminosarum* bv. *viciae* into root hairs requires minimal nod factor specificity, but subsequent infection thread growth requires nodO or nodE. *Mol. Plant-Microbe Interact.* 13:754–762
- Wang, Y.-H., Qu, X.-J., Chen, S.-Y., Li, D.-Z., and Yi, T.-S. 2017. Plastomes of Mimosoideae: structural and size variation, sequence divergence, and phylogenetic implication. *Tree Genet. Genomes.* 13:41
- Werner, G. D. A., Cornwell, W. K., Sprent, J. I., Kattge, J., and Kiers, E. T. 2014. A single evolutionary innovation drives the deep evolution of symbiotic N₂-fixation in angiosperms. *Nat. Commun.* 5:4087
- Wielgoss, S., Barrick, J. E., Tenaillon, O., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., Lenski, R. E., and Schneider, D. 2011. Mutation Rate Inferred From Synonymous Substitutions in a Long-Term Evolution Experiment With *Escherichia coli*. *G3• Genes|Genomes|Genetics.* 1:183–186
- Xiao, T. T., Schilderink, S., Moling, S., Deinum, E. E., Kondorosi, E., Franssen, H., Kulikova, O., Niebel,

- A., and Bisseling, T. 2014. Fate map of *Medicago truncatula* root nodules. *Development*. 141:3517–28
- Yano, K., Yoshida, S., Muller, J., Singh, S., Banba, M., Vickers, K., Markmann, K., White, C., Schuller, B., Sato, S., Asamizu, E., Tabata, S., Murooka, Y., Perry, J., Wang, T. L., Kawaguchi, M., Imaizumi-Anraku, H., Hayashi, M., and Parniske, M. 2008. CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proc. Natl. Acad. Sci.* 105:20540–20545
- Young, N. D., Debelle, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., Benedito, V. A., Mayer, K. F. X., Gouzy, J., Schoof, H., Van de Peer, Y., Proost, S., Cook, D. R., Meyers, B. C., Spannagl, M., Cheung, F., De Mita, S., Krishnakumar, V., Gundlach, H., Zhou, S., Mudge, J., Bharti, A. K., Murray, J. D., Naoumkina, M. A., Rosen, B., Silverstein, K. A. T., Tang, H., Rombauts, S., Zhao, P. X., Zhou, P., Barbe, V., Bardou, P., Bechner, M., Bellec, A., Berger, A., Bergès, H., Bidwell, S., Bisseling, T., Choisne, N., Couloux, A., Denny, R., Deshpande, S., Dai, X., Doyle, J. J., Dudez, A.-M., Farmer, A. D., Fouteau, S., Franken, C., Gibelin, C., Gish, J., Goldstein, S., González, A. J., Green, P. J., Hallab, A., Hartog, M., Hua, A., Humphray, S. J., Jeong, D.-H., Jing, Y., Jöcker, A., Kenton, S. M., Kim, D.-J., Klee, K., Lai, H., Lang, C., Lin, S., Macmil, S. L., Magdelenat, G., Matthews, L., McCarrison, J., Monaghan, E. L., Mun, J.-H., Najar, F. Z., Nicholson, C., Noirot, C., O’Bleness, M., Paule, C. R., Poulain, J., Prion, F., Qin, B., Qu, C., Retzel, E. F., Riddle, C., Sallet, E., Samain, S., Samson, N., Sanders, I., Saurat, O., Scarpelli, C., Schiex, T., Segurens, B., Severin, A. J., Sherrier, D. J., Shi, R., Sims, S., Singer, S. R., Sinharoy, S., Sterck, L., Viollet, A., Wang, B.-B., Wang, K., Wang, M., Wang, X., Warfsmann, J., Weissenbach, J., White, D. D., White, J. D., Wiley, G. B., Wincker, P., Xing, Y., Yang, L., Yao, Z., Ying, F., Zhai, J., Zhou, L., Zuber, A., Dénarié, J., Dixon, R. A., May, G. D., Schwartz, D. C., Rogers, J., Quétier, F., Town, C. D., and Roe, B. A. 2011. The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature*. 480:520–524
- Zhan, J., and Sun, Q. 2012. Diversity of free-living nitrogen-fixing microorganisms in the rhizosphere and non-rhizosphere of pioneer plants growing on wastelands of copper mine tailings. *Microbiol. Res.* 167:157–165
- Zhou, D., Udpa, N., Gersten, M., Visk, D. W., Bashir, A., Xue, J., Frazer, K. A., Posakony, J. W., Subramaniam, S., Bafna, V., and Haddad, G. G. 2011. Experimental selection of hypoxia-tolerant *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 108:2349–54