



HAL
open science

Study of terminal bacteroid differentiation features during the legume-rhizobium symbiosis

Quentin Nicoud

► **To cite this version:**

Quentin Nicoud. Study of terminal bacteroid differentiation features during the legume-rhizobium symbiosis. Plant breeding. Université Paris-Saclay, 2021. English. NNT: 2021UPASB007. tel-03208429

HAL Id: tel-03208429

<https://theses.hal.science/tel-03208429>

Submitted on 26 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Study of terminal bacteroid differentiation features during the legume-rhizobium symbiosis.

Étude de fonctions bactériennes impliquées dans la différenciation bactérienne lors de la symbiose rhizobium-légumineuse

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n°567, Sciences du végétal : du gène à l'écosystème (SEVE)
Spécialité de doctorat : Biologie
Unité de recherche : Université Paris-Saclay, CEA, CNRS,
Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France.
Référent : Faculté des Sciences d'Orsay

**Thèse présentée et soutenue à Paris-Saclay,
le 3 février 2021, par**

Quentin NICOUD

Composition du Jury

Marianne DELARUE

Professeure des Universités, Université Paris-Saclay (IPS2)

Présidente

Gabriella PESSI

Directrice de Recherche, Université de Zurich

Rapporteur & Examinatrice

Pierre FRENDU

Professeur des Universités, Université de Nice Sophia Antipolis (ISA)

Rapporteur & Examineur

Pierre CZERNIC

Professeur des Universités, Université de Montpellier (IPME)

Examineur

Direction de la thèse

Benoît ALUNNI

Maître de Conférences, HDR, Université Paris-Saclay (I2BC)

Directeur de thèse

Peter MERGAERT

Directeur de Recherche, CNRS (I2BC)

Co-encadrant et Examineur

ACKNOWLEDGEMENTS / REMERCIEMENTS

Je tiens ici à remercier toutes les personnes qui, de près ou de loin, par leur soutien humain ou leurs compétences scientifiques, ont participé à l'accomplissement de mon travail de thèse.

Tout d'abord, je tiens à remercier mes directeurs de thèse, Benoît Alunni et Peter Mergaert, qui m'ont guidé grâce à leurs précieux conseils, leur soutien, leur patience et leurs enseignements durant ces trois années de thèse. Vous avez toujours été à l'écoute et c'est grâce à toutes ces choses, que ces lignes sont aujourd'hui encrées ici. Particulièrement, je tiens à te remercier, Benoît, pour tes enseignements à l'université qui m'ont donnée goût à la biologie végétale puis aux interactions plantes-bactéries. Je tiens également à te remercier Peter pour ce premier stage dans ton équipe en M1 qui furent mes premiers pas en microbiologie et au sein de l'équipe. Toutes ces choses ont largement participé à construire ce que je suis aujourd'hui.

Également, je tiens à remercier les membres de mon jury de thèse : Marianne Delarue, Pierre Czernic, et en particulier Gabriella Pessi et Pierre Frendo pour avoir accepté d'évaluer mon travail. Je souhaite également remercier les membres de mon comité de thèse : Emanuele Biondi, Nicolas Mirouze, Souphian Ouchane et George C. diCenzo. Merci à vous tous pour votre aide et vos précieux conseils.

Je souhaite aussi remercier les collaborateurs d'Italie qui m'ont accueilli chaleureusement dans ma quête du savoir relatif aux modèles métaboliques, au cours d'un séjour d'un mois au sein de leur laboratoire à l'Université de *Sesto di Fiorentino*. Ainsi je remercie toute l'équipe COMBO et plus particulièrement le chef d'équipe Marco Fondi, Alessio Mengoni ainsi que George DiCenzo pour leur enseignement, accueil et immersion dans la culture Italienne. Ce séjour n'aurait pas été possible sans le financement prodigué par le LabEx SPS. Je remercie donc les membres du comité qui ont accepté mon dossier. Je souhaite également remercier Kar-Wai Hong pour son accueil lors de ma visite à l'Université de Jiangsu en Chine, ayant été un traducteur indispensable ainsi qu'un puit d'information concernant la culture asiatique. Également, merci pour ton partage de connaissances en matière de techniques de séquençage et pour ta bonne humeur. J'en profite également pour remercier le ministère de l'Éducation et de la Recherche pour le financement de ma bourse de thèse.

Evidemment, je tiens à remercier tous les membres permanents de mon équipe : Denis Faure, Yves Dessaux, Emanuele Biondi, Tania Timtchenko, Catherine Grandclément et Raynald Cossard. Plus particulièrement, merci Tania pour ta participation à mon encadrement en matière de biologie moléculaire ainsi que Catherine pour ta bonne humeur et ton soutien. Également, je remercie Emanuele pour avoir participé à l'encadrement au début de ma thèse sur l'étude du cycle cellulaire bactérien et son implication dans la différenciation (non présenté ici), avant le début du projet appliqué au

laboratoire. Merci à toi Corine pour toute l'aide que tu m'as apportée durant les expériences *in planta* notamment pour les dénudulations, rendant ces moments bien plus agréables. Bien sûr, j'ai également une pensée particulière pour toi Raynald, merci pour ta bonne humeur, ton humour et ton soutien.

Je souhaite également remercier tous les autres membres de l'équipe : Jérémy Cigna, Besma Bouznif, Etienne Jeanne, Gaëlle Lextrait, Romain Jouan, Abdillah Mohamed et Sara Dendene. Aux nouveaux thésards, ce fut bref mais intense, je vous souhaite tout le meilleur pour la suite de vos thèses. Je remercie aussi tous les « anciens »: Joy Lachat, Kevin Robic, Marta Torres Béjar, Tsubasa Ohbayashi, Pauline Blin, Fabienne Pierre et Almudena Gonzales-Mula avec lesquels j'ai partagé de bons moments. J'adresserai un remerciement particulier à Quentin Barrière et Florian Lamouche pour tous ces excellents moments passés au labo et au dehors. Merci à Nicolas Busset également pour les bons moments mais aussi pour ton soutien, notamment dans l'encadrement de stagiaires... Enfin je tiens à remercier Gregory Deicsics qui bien que n'ayant pas travaillé au sein de l'équipe a été en quelque sorte mon « binôme de thèse ». Merci pour ton soutien et pour ces excellents moments passés au labo et au dehors.

Ayant eu la chance de pouvoir enseigner à l'université dans le cadre d'un monitorat, je souhaiterais remercier tous les enseignants qui ont pu me conseiller durant cette aventure.

Enfin, je souhaite remercier toute ma famille, mon frère, mes parents, ma compagne et mes amis Paul, Anice, Anis, Marin, Maxime, Florence, Sophie, Eva et Léo pour m'avoir encouragé tout au long de ce travail, avoir cru en moi jusqu'au bout ou seulement m'avoir aidé à me changer les idées. Pour tous ces moments qui donnent à l'esprit une bouffée d'air frais permettant de rester opérationnel quand le mental est mis à l'épreuve, je vous remercie.

ABBREVIATIONS

A

ABC: ATP-binding Cassette
ADP: Adenosine Diphosphate
AHL: N-acyll Homoserine Lactone
AMP: Antimicrobial Peptides
AS: Asparagine Synthase
ATP: Adenosyl Triphosphate

B

bv: biovar

C

CCaMK: Ca²⁺/Calmodulin-dependent
protein Kinase
CDK: Cyclin-deependent Kinase
Cfu: colony forming units
CG: Cyclic Betaglucans
CoBRA: Constraint Based Reconstruction
and Analysis
CPS: Capsular Polysaccharides
CSSP: Common Symbiosis ymbiosis
Signaling Pathway

D

DAP: Differentially Accumulated Proteins
DEG: Differentially Expressed Genes
DNA: Deoxyribonucleic Acid
dnf: defective in nitrogen fixation
dpi: days post-inoculation
DRM: Detergent Resistant Membranes

E

EPS: Exo Polysaccharides
ER: Endoplasmic Reticulum
ESR: Extra-cytoplasmic/Envelope Stress
Response

F

FBA: Flux Balance Analysis

G

GFP: Green Fluorescent Protein
GRN: Gene Regulation Networks
GRP: Glycine Rich Peptides
GS: Glutamine Synthase

H

HGT: Horizontal Gene Transfer
HITS: High throughput Insertion Track by
deep Sequencing
HR: Homologous Recombination

I

ICE: Integrative and Conjugative Elements
IM: Innner Membrane
INSeq: Insertion Sequencing

IRLC: Inverted-Repeat-Lacking Clade

K

KDO: 2-keto-3deoxyoctonic acid
KPS: K-antigen Polysaccharides

L

LCO: Lipo-Chito-Oligosaccharide
LPS: Lipo-Polysaccharides
LR: Lipid Raft
LRR: Leucine-Rich-Repeats

M

MM: Metabolic Model
MIC: Minimum Inhibitory Concentration

N

NBD: Nucleotide Binding Domain
NCR: Nodule-specific Cysteine-Rich
peptides

O

OM: Outer Membrane

P

PBP: Periplasmic-binding Protein
PHB: Poly-Hydroxy-Butyrate
PI: Propidium Iodide
Pi: Phosphate Inorganic

Q

QS: Quorum Sensing

R

RLK: Receptor-Like Kinase
ROS: Reactive Oxygen Species
rRNA: ribosomal Ribonucleic Acid

S

sp.: species
ssp.: subspecies

T

TBD: Terminal Bacteroid Differentiation
TCA: Tricarboxylic Acid
TEM: Transmission Electron Microscopy
TIS: Transposon Insertion Sequencing
Tn-Seq: Transposon Sequencing
TraDIS: Transposon-Directed Insertion site
Sequencing
T3SS: Type 3 Secretion System

U

UV: Ultra-Violet

V

var: variety
VLCFA: Very Long Chain Fatty Acid

W

WT: Wild-Type

LIST OF ORGANISM ABBREVIATIONS

Abbreviation	Full name	Short description
<i>A. afraspera</i>	<i>Aeschynomene afraspera</i>	Dalbergioid legume
<i>A. evenia</i>	<i>Aeschynomene evenia</i>	Dalbergioid legume
<i>A. indica</i>	<i>Aeschynomene indica</i>	Dalbergioid legume
<i>B. aphidicola</i>	<i>Buchnera aphidicola</i>	insect symbiont
<i>B. diazoefficiens</i>	<i>Bradyrhizobium diazoefficiens</i>	plant symbiont
<i>B. elkanii</i>	<i>Bradyrhizobium elkanii</i>	plant symbiont
<i>E. coli</i>	<i>Escherichia coli</i>	model enterobacterium
<i>E. scolopes</i>	<i>Euprynma scolopes</i>	bobtail squid
<i>G. max</i>	<i>Glycine max</i>	soybean (legume)
<i>L. japonicus</i>	<i>Lotus japonicus</i>	trefoil (legume)
<i>M. loti</i>	<i>Mesorhizobium loti</i>	plant symbiont
<i>M. sativa</i>	<i>Medicago sativa</i>	alfalfa (IRLC legume)
<i>M. truncatula</i>	<i>Medicago truncatula</i>	alfalfa (IRLC legume)
<i>P. sativum</i>	<i>Pisum sativum</i>	pea (IRLC legume)
<i>P. vulgaris</i>	<i>Phaseolus vulgaris</i>	bean (legume)
<i>S. enterica</i>	<i>Salmonella enterica</i>	human pathogen
<i>S. fredii</i>	<i>Sinorhizobium fredii</i>	plant symbiont
<i>S. meliloti</i>	<i>Sinorhizobium meliloti</i>	plant symbiont
<i>V. fischerii</i>	<i>Vibrio fischerii</i>	squid symbiont
<i>V. sativa</i>	<i>Vicia sativa</i>	vetch (IRLC legume)

RESUMÉ EN FRANÇAIS

i. Contexte général

La disponibilité de l'azote est une limitation majeure pour le développement des plantes dans de nombreux environnements, y compris les milieux agricoles. Pour surmonter ce problème et prospérer sur des substrats présentant une faible teneur en azote, les cultures sont fortement fertilisées, ce qui entraîne d'importants dommages environnementaux et des inconvénients financiers (Erisman et al. 2016 ; Zhao et al., 2017). Les plantes de la famille des légumineuses ont acquis la capacité de former des associations symbiotiques avec les bactéries du sol, les rhizobia, qui fixent l'azote atmosphérique au profit des plantes. La symbiose légumineuse - rhizobium est un processus écologique majeur dans le cycle de l'azote, responsable du principal apport d'azote fixe dans la biosphère et peut fournir l'azote nécessaire à la croissance des plantes en agriculture. L'efficacité de cette symbiose repose sur la coévolution des partenaires. Ces associations symbiotiques conduisent au développement de nodules - des organes racinaires qui abritent des rhizobia. Dans ces nodules, les rhizobia adoptent un mode de vie intracellulaire et se différencient en bactéroïdes qui convertissent l'azote atmosphérique en ammoniac, avant de le transférer à la plante. Des étapes critiques de reconnaissance ont lieu tout au long du processus symbiotique et définissent la compatibilité des partenaires végétaux et bactériens (Oldroyd, 2013). Si les mécanismes impliqués dans les premiers stades de la symbiose sont bien décrits, ceux des stades ultérieurs sont beaucoup moins clairs et pourraient affecter non seulement la capacité d'interaction mais aussi l'efficacité de la symbiose (c'est-à-dire l'avantage pour la plante).

Selon la combinaison de plantes hôtes et de souches rhizobiennes, le résultat des interactions rhizobium-légumineuses varie d'associations non fixatrices à des symbioses très bénéfiques pour la plante. Certaines légumineuses, mais pas toutes, optimisent leur retour sur investissement dans la symbiose en imposant à leurs micro-symbiote un programme de différenciation terminale qui augmente leur efficacité symbiotique mais impose un niveau de stress élevé et réduit drastiquement leur viabilité. Il a été proposé que les peptides antimicrobiens riches en cystéines (NCR) spécifiques aux nodules, produits par les légumineuses du groupe des dalbergioïdes et le Clade à répétition inversée manquante (IRLC), jouent un rôle crucial dans le contrôle de la spécificité de l'hôte-symbiote au stade intracellulaire de la symbiose (Gourion & Alunni, 2018). Les peptides NCR ciblent les

bactéroïdes et régissent leur différenciation (Mergaert et al., 2003 ; Mergaert et al., 2006). Chez les légumineuses susmentionnées, le processus de différenciation entraîne des changements si profonds qu'ils suppriment la capacité des bactéroïdes à reprendre leur croissance et est donc appelé différenciation bactérienne terminale (TBD). La TBD se démarque de la formation de bactéroïdes chez les légumineuses dépourvues de gènes NCR (par *exemple le soja*), où les bactéroïdes sont dans un état réversible et peuvent reprendre leur croissance lorsqu'ils sont libérés des nodules (Alunni & Gourion, 2016). Plus précisément, la TBD est associée à l'élongation cellulaire, une augmentation du contenu en ADN bactérien par un changement de cycle cellulaire vers l'endo-réplication (Mergaert et al., 2006 ; Czernic et al., 2015). En outre, une perméabilité accrue de l'enveloppe bactérienne se développe au cours de la TBD, très probablement en raison de l'interaction des peptides NCR avec les membranes bactériennes (Mergaert et al., 2006 ; van de Velde et al., 2010 ; Farkas et al., 2014). Ensemble, ces altérations de la physiologie bactérienne sont associées à une forte diminution de la viabilité des bactéries différenciées, qui ne parviennent pas à retrouver leur croissance lorsqu'elles sont extraites des nodules (Mergaert et al., 2006).

Les peptides NCR ont des caractéristiques communes avec les peptides antimicrobiens (AMP) et au moins certains NCR, en particulier les cationiques, peuvent tuer ou inhiber la croissance *in vitro* non seulement des symbiotes du rhizobium mais aussi de nombreuses autres bactéries et même de champignons (Lima et al., 2020). Leur activité antibactérienne majeure résulte de leur capacité à perturber l'intégrité des membranes internes et externes des bactéries, entraînant une perte du potentiel membranaire (Mergaert et al., 2003). En outre, les peptides NCR ont des cibles intracellulaires, se liant aux ribosomes et inhibant la traduction (Farkas et al., 2004). Cependant, l'élimination des endosymbiontes n'est évidemment pas ce qui se passe dans les cellules de nodules symbiotiques et les bactéroïdes restent très longtemps actifs malgré la charge élevée des NCR. Il est possible que l'environnement des cellules de nodules symbiotiques et des symbiosomes contribue à tempérer l'activité antimicrobienne des peptides. Il est important de noter que les fonctions spécifiques des bactéries elles-mêmes sont également des déterminants de la résistance aux NCR dans les bactéroïdes.

La résistance aux AMP est essentielle chez les agents pathogènes pour surmonter la réponse immunitaire innée et établir une infection chronique de l'hôte. Les mutants pathogènes présentant une résistance plus ou moins élevée aux AMP perdent ou gagnent en pathogénicité, respectivement (Mergaert et al., 2018). Les commensales de l'intestin ont besoin de fonctions de résistance aux AMP pour assurer la résilience dans l'environnement changeant de l'intestin (Cullen et al., 2015). De même, *Sinorhizobium meliloti*, le symbiote des plantes *Medicago*, a besoin de protéines de résistance aux AMP pour l'infection chronique des cellules de nodules. BacA (connu sous le nom de BclA chez les bradyrhizobia qui infecte les *Aeschynomene* producteurs de NCR (Guefrachi et al., 2015 ; Barrière et al., 2017) est un transporteur de peptides nécessaire pour contrer les peptides NCR à l'intérieur des cellules nodulaires symbiotiques (Haag et al., 2011). Les mutants de *S. meliloti bacA* sont hypersensibles aux peptides NCR antimicrobiens. Ils induisent des nodules et infectent leurs cellules symbiotiques de façon apparemment normale, mais les mutants meurent dès qu'ils sont libérés dans les cellules symbiotiques. Cette mort peut être évitée en bloquant le transport des NCR vers le rhizobium infectant chez le mutant *M. truncatula dnf1* (Haag et al., 2011). Les protéines BacA et BclA sont des transporteurs bactériens de peptides avec un large spectre d'activité d'importation de peptides (Guefrachi et al., 2015 ; Barrière et al., 2017). Elles peuvent favoriser l'absorption des peptides NCR, ce qui suggère que le transport des peptides NCR à travers les BacA ou BclA crée une résistance en les redirigeant loin de la membrane bactérienne, limitant ainsi les dommages à la membrane. Les exopolysaccharides (EPS) sont un autre facteur connu de *S. meliloti* qui aide les endosymbiotes à résister aux NCR (Arnold et al., 2018). Ce polysaccharide extracellulaire chargé négativement piège les AMP cationiques, réduisant ainsi leur concentration effective au voisinage de la membrane. Enfin, certains rhizobia possèdent une protéase dégradant les NCR qui leur permet d'échapper totalement au processus de différenciation des bactéries et qui augmente leur prolifération dans les nodules sans fixer l'azote, ce qui entraîne un mode de vie plus parasitaire de ces rhizobia dans les nodules (Price et al., 2015).

La résistance bactérienne aux AMP est généralement multifactorielle (Cole & Nizet, 2016), ce qui suggère qu'outre le BacA et l'EPS, des fonctions supplémentaires des bactéroïdes *S. meliloti* contribuent à la résistance des NCR dans les cellules de nodules symbiotiques. La littérature sur *S. meliloti* est riche en description des gènes bactériens nécessaires à la

symbiose. Cependant, les rapports sur ces mutants manquent souvent d'informations précises sur leur phénotype bactérien et/ou sur leur sensibilité aux NCR. En outre, les analyses du transcriptome et du Tn-seq (séquençage des transposons) des cellules traitées par les NCR et les études d'interaction NCR-protéine ont permis d'identifier toute une série de fonctions candidates supplémentaires répondant aux NCR chez *S. meliloti* (Farkas et al., 2014 ; Arnold et al., 2017 ; Penterman et al., 2014). Avec BacA et les EPS, certaines de ces fonctions de *S. meliloti* pourraient contribuer à atténuer le stress des NCR sur les bactéroïdes. Pour tester cette hypothèse, nous avons sélectionné dans la première étude trois fonctions candidates et analysé le phénotype des mutants correspondants dans la résistance NCR et la formation des bactéroïdes.

Bradyrhizobium diazoefficiens USDA110 a été isolé en tant que symbiote du soja mais il peut également établir une interaction symbiotique fonctionnelle avec *Aeschynomene afraspera*. Contrairement au soja, *A. afraspera* déclenche une différenciation bactérienne terminale, un processus impliquant l'élongation des cellules bactériennes, la polypléidie et la perméabilité de la membrane, ce qui entraîne une perte de viabilité bactérienne tandis que les plantes augmentent leur bénéfice symbiotique. Dans une seconde étude, une combinaison de métabolomique végétale, de protéomique bactérienne et de transcriptomique, ainsi que des analyses cytologiques ont été utilisées pour étudier la physiologie des bactéroïdes USDA110 dans ces deux plantes hôtes. Nous avons montré que l'USDA110 établit une symbiose peu efficace avec *A. afraspera*, malgré l'activation complète du programme symbiotique bactérien. Nous avons trouvé des signatures moléculaires de niveaux de stress élevés chez les symbiotes d'*A. afraspera*, alors que celles de la différenciation bactérienne terminale n'étaient que partiellement activées. Enfin, nous avons montré que chez *A. afraspera*, les bactéroïdes USDA110 subissent une différenciation terminale atypique marquée par la déconnexion des caractéristiques canoniques de ce processus. Cette étude montre comment une souche de rhizobium peut adapter sa physiologie à un nouvel hôte et faire face à la différenciation terminale alors qu'elle n'a pas co-évolué avec un tel hôte.

ii. Etude des fonctions de Sinorhizobium meliloti requises pour la résistance aux peptides NCR et pour la différenciation bactérienne.

Les fonctions de *S. meliloti* sélectionnées dans cette étude comprennent un transporteur d'absorption peptidique à large spécificité codé par les gènes *yjABEF* (*SMc02829-SMc02832*), des modifications de lipopolysaccharides (LPS) médiées par *lpsB* (*SMc01219*) et *lpxXL* (*SMc04268*) ainsi que *rpoH1* (*SMc00646*), codant un facteur sigma de stress. Le transporteur YejABEF ABC, composé de la protéine de liaison périplasmique YejA (*SMc02832*), de l'ATPase YejF (*SMc02829*) et des perméases YejB (*SMc02831*) et YejE (*SMc02830*), a été sélectionné sur la base du fait que ce transporteur et son orthologue *Escherichia coli* ont une spécificité d'absorption peptidique chevauchante avec le BacA (connu sous le nom de SbmA dans *E. coli* (Bantysh et al., 2015). En outre, les mutants dans les gènes orthologues des agents pathogènes animaux *Salmonella enterica* serovar *Typhimurium* et *Brucella melitensis* sont plus sensibles à divers peptides antimicrobiens, ont une virulence moindre et une capacité réduite à persister dans leur hôte animal (Eswarappa et al., 2008 ; Wang et al., 2016). Enfin, et surtout, un dépistage génétique par Tn-seq chez *S. meliloti* a révélé que le transporteur mutant a une sensibilité accrue au peptide NCR247 (Arnold et al., 2017). Les AMP se lient au LPS anionique dans la phase initiale d'interaction avec la membrane bactérienne et s'insèrent ensuite dans la couche lipidique composée de la fraction lipidique A du LPS. Par conséquent, la structure des LPS est l'un des principaux déterminants de la résistance et de la sensibilité aux AMP chez les bactéries Gram-négatives (Cole & Nizet, 2016). Les gènes sélectionnés *lpsB* et *lpxXL* codent respectivement une glycosyltransférase impliquée dans la synthèse du noyau du LPS et une acyl-transférase d'acide gras à très longue chaîne impliquée dans la biosynthèse du lipide A. Les mutants de ces gènes sont affectés dans la résistance aux AMP et dans la symbiose (Campbell et al., 2003 ; Haag et al., 2009). Enfin, le gène *rpoH1* est un régulateur de stress global chez *S. meliloti* (Barnett et al., 2012). Ce gène ainsi que ses cibles sont régulés à la hausse dans les cellules traitées par NCR247 (Penterman et al., 2014). Un mutant *rpoH1* est également affecté en symbiose (Mitsui et al., 2004). Ces gènes sélectionnés sont exprimés dans des nodules, avec un pic d'expression dans différentes régions du nodule où les bactéries infectent les cellules végétales, subissent le processus de différenciation ou fixent l'azote.

Les mutants de ces gènes candidats ont été construits (*yejA*, *yejE*, *yejF*, *rpoH1*) ou ont été obtenus à partir d'études antérieures (*lpsB*, *lpxXL*). Les bactéries en phase logarithmique ont été traitées avec un petit panel de peptides NCR dont l'activité antimicrobienne avait été précédemment démontrée (Montiel et al., 2017). Les mutants analysés dans cette étude ont tous montré une plus grande sensibilité à au moins un des peptides par rapport au type sauvage.

Ensuite, nous avons testé si les gènes sélectionnés de *S. meliloti* sont nécessaires à l'établissement d'une symbiose fonctionnelle avec *M. truncatula*. Le phénotype symbiotique de ces mutants a été comparé à la souche de type sauvage et au mutant *bacA*. L'inspection macroscopique du système racinaire des plantes inoculées avec la souche sauvage et sept mutants a révélé que les mutants des gènes *yejA*, *yejE*, *yejF* et *lpxXL* formaient des nodules d'aspect similaire aux nodules fonctionnels formés par la souche de type sauvage. D'autre part, les mutants *bacA*, *lpsB* et *rpoH* ont formé des nodules d'aspect anormal, petits et blancs, symptomatiques de l'absence de fixation de l'azote (Fix-) et en accord avec les descriptions précédentes (Haag et al., 2011 ; Campbell et al., 2003 ; Mitsui et al., 2004). L'activité de fixation de l'azote des nodules a été mesurée directement avec le test de réduction de l'acétylène (ARA). Comme prévu par les observations macroscopiques, les nodules infectés par les mutants *bacA*, *lpsB* et *rpoH1* étaient non fonctionnels (Fix-). Les autres mutants, dans les gènes *yejA*, *yejE*, *yejF* et *lpxXL*, présentaient une activité de fixation de l'azote (Fix+), toujours en accord avec les caractéristiques visuelles des nodules. Néanmoins, les mutants *yejE*, *yejF* et *lpxXL* avaient une activité réduite par rapport au type sauvage.

L'organisation histologique des nodules formés par les mutants, et en particulier la formation de cellules symbiotiques infectées et la viabilité des bactéries qu'elles contiennent, a été analysée. Etonnement, la morphologie des bactéroïdes mutants *yejA*, *yejE* et *yejF* a été sensiblement modifiée par rapport aux bactéroïdes de type sauvage et présentent pour une part d'entre eux une forte perméabilisation de leurs membranes. L'élargissement des bactéroïdes a été confirmé par microscopie électronique à transmission et par cytométrie en flux. Cette dernière technique a également révélé que ce changement de morphologie est associé à une augmentation de la quantité d'ADN dans les cellules. La microscopie électronique à transmission a également révélé que le mutant *yejF* présente une désorganisation de son

enveloppe et accumule des marqueurs de stress. Le mutant *lpxXL* présentait des bactéroïdes allongés mais surtout très perméabilisés. Confirmant l'importance de ce gène pour la survie des bactéroïdes à la NCR chez *M. truncatula*, sans qu'il soit crucial pour l'infection et la fixation de l'azote. À l'intérieur des initiations de nodules provoquées par le mutant *lpsB*, aucune cellule ne semble être colonisée par les bactéries comme le révèle la microscopie confocale. Ce gène semble être requis à un stade plus précoce de l'infection. Le niveau de différenciation observé a été confirmé grâce à des expériences de cytométrie en flux permettant d'estimer la taille et la quantité d'ADN des cellules. L'utilisation d'un mutant de *M. truncatula* ne sécrétant pas de peptide NCR, a confirmé que la forte perméabilisation observée chez certains mutants était liée aux NCRs. Le mutant *lpsB* n'a pas formé de structures nodulaires détectables sur les racines des plants mutants. Nous ne pouvons donc pas conclure sur l'implication des peptides NCR dans le phénotype symbiotique de ce mutant.

La forte anomalie bactérienne observée chez les mutants du transporteur *yej* est *a priori* contradictoire avec l'activité de fixation de l'azote de ces nodules. Pour analyser spécifiquement l'activité de ces bactéries aberrantes dans les nodules, nous avons analysé l'activité du promoteur *nifH* dans ces bactéroïdes. Le gène *nifH* code la sous-unité de l'enzyme nitrogénase et son expression est un marqueur mettant en évidence les bactéroïdes fixateurs d'azote. Un gène GFP sous le contrôle du promoteur *nifH* situé sur le plasmide *ppNifH-GFP* a été introduit dans la souche de type sauvage et les mutants *bacA* (comme témoin négatif), *yejE* et *yejF*. Nous avons montré que dans les nodules infectés par les mutants *yejE* et *yejF* équipés de la construction rapportrice *pnifH*:GFP les cellules symbiotiques pouvaient contenir uniquement des bactéries exprimant le gène *nifH*, uniquement des bactéries négatives pour le gène *nifH* et perméabilisées ou un mélange des deux. Néanmoins, la microscopie a clairement montré que les bactéroïdes déformés induits chez ces mutants peuvent exprimer fortement *nifH* et donc probablement fixer l'azote au moins temporairement. Comme prévu dans le mutant *bacA*, aucun signal de GFP n'a été détecté, tandis que la souche sauvage toutes les cellules présentaient un signal de GFP. Ces résultats ont été confirmés par cytométrie en flux.

Chez *E. coli*, YejABEF et SbmA (BacA) servent de médiateurs pour le transport des peptides tels que la microcine C. Nous avons donc testé si le chevauchement des substrats de YejABEF et BacA peut être étendu à d'autres substrats SbmA/BacA connus. La bléomycine est

un peptide non ribosomique dont l'activité endommage l'ADN et qui doit donc être internalisé. Elle est médiée par la SbmA dans *E. coli* et le BacA dans *S. meliloti* (46). Nous avons estimé la sensibilité de nos souches à ce peptide, donnant ainsi une évaluation indirecte de leur capacité à transporter le peptide. Alors que le sauvage est sensible à la bléomycine et que le mutant *bacA* y est résistant, les trois mutants *yej* ont présenté une résistance significativement accrue comparée aux deux souches témoin. Un autre peptide antimicrobien a été testé, nommée Bac7 (son nom n'est pas lié au BacA ou aux bactéroïdes). Il est d'origine animale, inhibe l'activité des ribosomes et est importé via SbmA ou BacA (Marlow et al., 2009 ; Mardirossian et al., 2014). Grâce à un test basé sur la cytométrie de flux et un dérivé fluorescent du peptide Bac7, nous avons constaté que les trois mutants *yej* n'étaient pas affectés dans l'internalisation du Bac7, excluant l'implication du transporteur YejABEF dans l'absorption de ce peptide, contrairement à BacA. Enfin, le peptide NCR247 est un autre substrat connu du BacA (Barrière et al., 2017 ; Guefrachi et al., 2015). Nous avons testé l'impact du transporteur YejABEF sur l'absorption de NCR247 avec un test similaire à celui du test Bac7 ci-dessus. L'analyse a montré que si l'absorption de NCR247 est complètement supprimée chez le mutant *BacA*, comme prévu, et comme indiqué précédemment, son absorption est également réduite mais pas complètement supprimée chez les trois mutants *Yej*. Cela suggère que le transporteur contribue à l'absorption du NCR.

iii. La nodulation d'Aeschynomene afraspera par Bradyrhizobium diazoefficiens USDA110 est associée à une différenciation terminale atypique des bactéroïdes ainsi qu'à une efficacité symbiotique suboptimale.

B. diazoefficiens USDA110, le symbiote modèle du soja, est capable d'établir une symbiose fonctionnelle avec *A. afraspera*. Cette dernière est un hôte phylogénétiquement distant appartenant au clade des Dalbergioïdes, qui interagit naturellement avec des rhizobia photosynthétiques telles que *Bradyrhizobium* sp. ORS285 (Renier et al., 2011 ; Ledermann et al., 2018). Grâce à des mesures de l'activité de la nitrogénase, ainsi que des teneurs en azote et en carbone nous avons montré qu'USDA110 semble fixer l'azote moins efficacement qu'ORS285 chez *A. afraspera*. L'évaluation de la biomasse des plantes infectées par ces deux souches confirment ces résultats puisque le rapport entre la masse de la partie aérienne et de

la partie racinaire reflétant l'état nutritionnel de la plante, est réduit chez les plantes *A. afraspera* nodulées par USDA110. Pour caractériser davantage cette symbiose sous-optimale, nous avons analysé le métabolome de nodules entiers. L'allantoïne, qui est connue pour être la principale forme d'azote exportée par les nodules de soja, y est spécifiquement détectée (Collier et al., 2012). Au contraire, l'asparagine et la glutamine sont les principaux composés azotés exportés dans les nodules d'*A. afraspera* et leur quantité est inférieure dans les nodules infectés par l'USDA110 par rapport aux nodules infectés par l'ORS285, ce qui indique une fixation réduite de l'azote par les bactéroïdes (Lamouche et al., 2019). En outre, une accumulation de saccharose, d'acide phosphorique et d'ascorbate, et à l'inverse, une forte réduction de la teneur en tréhalose est constatée dans les nodules d'*A. afraspera* infectés par USDA110 spécifiquement. L'accumulation de ces composés suggère un stress au niveau des nodosités ainsi qu'un déficit dans la fixation d'azote, reflétant le trouble métabolique apparent dans les nodules infectés par l'USDA110.

Afin de mieux comprendre l'interaction mal-adaptée entre USDA110 et *A. afraspera*, la physiologie des bactéries a été évaluée grâce à l'analyse du transcriptome et du protéome. Les bactéries efficaces du soja et les cellules USDA110 libres cultivées en milieu riche (phase de croissance exponentielle en condition aérobie) ont servi de référence. Dans le jeu de données sur les protéomes, 1808 protéines USDA110 ont été identifiées. Dans l'ensemble de données transcriptomiques, 3150 gènes sont différentiellement exprimés dans au moins une condition (gènes différentiellement exprimés ou DEG). Parmi les 1808 protéines identifiées, 815 présentent une accumulation différentielle (protéines accumulées différentiellement ou DAP) et 438 des gènes apparentés sont également exprimés différentiellement dans les données des transcriptomes, tandis que 175 DEG ne sont pas des DAP. Parmi les 815 DAP, 705 et 699 protéines sont accumulées de manière significativement différente dans *G. max* et *A. afraspera* respectivement par rapport à la culture bactérienne contrôle. Dans les ensembles de données transcriptomiques, 1999 DEG entre la culture bactérienne et les bactéroïdes, quel que soit l'hôte, ont été identifiés. Parmi eux, 1076 gènes présentaient une expression plus élevée dans les nodules. Un ensemble supplémentaire de 923 gènes a été réprimé *in planta*.

En limitant l'analyse aux fonctions bactériennes qui sont à la fois différentiellement exprimées (DEG) et différentiellement accumulées (DAP) dans les plantes chez les deux hôtes

par rapport à la culture bactérienne, nous avons identifié 222 gènes/protéines, dont 150 sont régulés à la hausse et 72 sont réprimés dans les plantes respectivement.

Les données sur le protéome et le transcriptome ont fourni une vue cohérente du métabolisme de fixation de l'azote de *B. diazoefficiens* dans les conditions testées. Les enzymes clés impliquées dans la respiration microaérobie et la fixation de l'azote ont été détectées parmi les protéines ayant le plus grand nombre de spectres dans les échantillons de nodules et les gènes correspondants sont parmi les plus fortement exprimés dans les bactéroïdes, tout en étant presque indétectables dans les conditions de vie libre. Il s'agit par exemple des sous-unités de la nitrogénase ainsi que des gènes nécessaires pour son assemblage. En plus de ces fonctions bactériennes attendues, de nombreuses autres protéines s'accumulent spécifiquement et fortement dans les deux types de nodules. C'est le cas des chaperonnes GroEL1/GroES1, du système d'absorption de l'hydrogène (gènes *hup*) ou encore de l'acide 1-aminocyclopropane-1-carboxylique (ACC) désaminase (*blr0241*). À l'inverse, les gènes de motilité codant pour les sous-unités des flagelles (*blI5844-blI5846*), les enzymes métaboliques et les sous-unités de transport sont fortement régulés à la baisse pendant la symbiose et sont difficilement détectables au niveau des protéines *in planta*. Prises ensemble, ces données montrent que les deux types de bactéries présentent un métabolisme typique orienté vers la fixation de l'azote, avec un arrêt partiel des fonctions ménagères. Cela indique que malgré la réduction apparente de l'efficacité symbiotique de l'USDA110 dans les nodules d'*A. afraspera*, la bactérie exprime pleinement son programme symbiotique au sein de cet hôte non indigène comme elle le fait dans le soja, son hôte d'origine.

La comparaison des mutants bactériens *A. afraspera* et *G. max* a également révélé des différences significatives dans les protéomes et les transcriptomes. Au niveau transcriptomique, 935 DEG ont pu être identifiés entre les deux types de bactéroïdes (509 *A. afraspera* > *G. max* et 426 *G. max* > *A. afraspera*). Cependant, si l'on considère uniquement les fonctions qui présentent des différences congruentes et significatives en termes de transcrits et de niveaux de protéines entre les plantes hôtes, on tombe à 63 gènes/protéines, 33 étant induits dans les nodules d'*A. afraspera* et 30 dans les nodules de *G. max*.

Il est intéressant de noter que la voie de dégradation de l'acide phénylacétique (PaaABCDEIK, *blr2891-blr2897*) a été fortement exprimée dans les nodules d'*A. afraspera* (bien

que seuls PaaABCD et PaaK aient été détectés par la protéomique), ainsi qu'un groupe de gènes non encore caractérisés qui seraient impliqués dans la dégradation du toluène (*blr3675-blr3680*). Le chaperon GroEL2 est également spécifiquement induit dans *A. afraspera*. De même, trois peptidases S1 (Dop : *blr2591*, *blr3130* et *blr7274*) sont fortement exprimées dans les nodules de ce dernier hôte avec une pompe d'efflux RND (*bll3903*) et une protéine à motif LTXXQ (*bll6433*), un motif que l'on retrouve également dans la réponse au stress périsplasmique CpxP26. La suraccumulation de ces protéines suggère que les bactéroïdes sont confrontés à des conditions de stress pendant cette interaction avec *A. afraspera*. Une protéine non caractérisée de liaison au soluté du transporteur ABC (*blr7922*) a également été surexprimée dans *A. afraspera*.

Une hydrolase de $\alpha\beta$ (*blr6576*) et une protéine semblable au récepteur TonB-dépendant (*bll2460*) ont été suraccumulées de manière spécifique à *G. max*. De même, un groupe métabolique non caractérisé comprenant des transkétolases (*blr2167-blr2170*), l'enzyme biosynthétique HemN1 (*bll2007*) et, dans une moindre mesure, une anthranilate phosphoribosyl-transférase (TrpD codée par *bll2049*) sont surexprimés dans les nodules de soja.

Dans une étude précédente (Lamouche et al., 2019), une analyse du transcriptome a été réalisée sur *Bradyrhizobium* sp. ORS285 en interaction avec *A. afraspera* et en culture. ORS285 est une souche qui a co-évolué avec *A. afraspera*, conduisant à une symbiose efficace caractérisée par une élongation cellulaire et une polyploïdisation des bactéroïdes. Afin de comparer l'expression génétique de ces deux rhizobia formant des nodules en culture et *in planta*, nous avons déterminé l'ensemble des gènes orthologues entre les deux souches. Cette analyse a donné un total de 3725 gènes. Seuls 343 de ces gènes ont présenté une expression différentielle (FDR < 0,01 et |LFC| > 1,58) entre la culture témoin et la condition *in planta*, et ce pour les deux bactéries. Une majorité de ces gènes (86,8 %) présentaient des profils d'expression congruents. Tout d'abord, les gènes *nif*, *fix* et *hup* sont couramment et fortement induits dans les deux souches au cours de leur vie symbiotique avec *A. afraspera*, ce qui est la marque d'une symbiose fonctionnelle.

En outre, 1026 gènes ont été exprimés de manière différentielle uniquement chez ORS285, et de même, il y avait 604 DEG spécifiques à USDA110. Par exemple, la voie sécrétoire

générale semble être spécifiquement induite dans ORS28517. A l'inverse, USDA110 présente une induction des gènes *rhcJQRU* qui sont impliqués dans l'injection de protéines effectrices de type 3 qui peuvent être importantes pour l'établissement de la symbiose alors qu'elles ne sont pas induites ou même réprimées chez ORS285.

Dans une description précédente de l'interaction entre *A. afraspera* - *B. diazoefficiens* USDA110, les caractéristiques typiques des TBD n'ont pas été observées et les bactéroïdes étaient très similaires à ceux de *G. max* où aucun TBD n'est présent (Barrière et al., 2017). Prises ensemble, les données moléculaires provenant des transcriptomes et protéomes n'indiquent pas clairement si les bactéroïdes USDA110 subissent une TBD chez *A. afraspera*. Nous avons donc étudié les caractéristiques des bactéroïdes USDA110 dans les nodules d'*A. afraspera* de manière plus détaillée.

Nous avons analysé les caractéristiques de différenciation des bactéroïdes dans les bactéroïdes USDA110 extraits des nodules de soja et d'*A. afraspera*. L'interaction entre *A. afraspera* et *Bradyrhizobium* sp. ORS285 a été utilisée comme témoin positif pour les caractéristiques à déterminer (Czernic et al., 2015 ; Bonaldi et al., 2011 ; Guefrachi et al., 2015). Le TBD est caractérisé par une élongation cellulaire. Nous avons quantifié la longueur, la largeur, la surface et la forme des cellules des bactéroïdes purifiés et des témoins de culture. Alors que les bactéroïdes ORS285 étaient plus volumineux dans les nodules d'*A. afraspera* que leurs homologues vivant en liberté, les bactéroïdes USDA110 étaient similaires aux bactéries vivant en liberté dans le soja et *A. afraspera*. Une autre caractéristique de la TBD est l'endoréplication. L'analyse de la teneur en ADN bactérien des bactéroïdes ORS285 dans *A. afraspera* par cytométrie de flux montre des pics à 6C et plus (Czernic et al., 2015). Comme prévu, les bactéroïdes USDA110 dans *G. max* ne donnent que deux pics, à 1C et 2C, de manière similaire aux cellules en cycle dans l'échantillon de culture bactérienne (Barrière et al., 2017). Des résultats similaires ont été obtenus pour USDA110 en interaction avec *A. afraspera*. Ainsi, en ce qui concerne la teneur en ADN et la taille des cellules, les bactéroïdes d'USDA110 ne présentent pas les caractéristiques typiques dans les nodules d'*A. afraspera*. La perte d'intégrité de la membrane est une troisième caractéristique des TBD qui contribue probablement fortement à la perte de viabilité des bactéroïdes. Une analyse de l'absorption de l'iodure de propidium (IP) par les bactéroïdes et les contrôles de culture correspondants ont été effectués

pour évaluer la perméabilité des bactéroïdes. Vingt minutes après l'application de l'IP, les bactéroïdes USDA110 d'*A. afraspera* présentent une perméabilité accrue qui est beaucoup plus proche des bactéroïdes ORS285 en interaction avec *A. afraspera* que de la faible perméabilité des bactéroïdes USDA110 des nodules *G. max*. Pour analyser la viabilité bactérienne, les bactéroïdes extraits des nodules ont été étalés sur milieu riche et les unités formatrices de colonies (cfu) ont été déterminées. Chez *G. max*, USDA110 a formé $1,46 \times 10^{10}$ colonies/mg de nodules (~100% de survie). A l'inverse, ORS285 n'a formé que $5,42 \times 10^7$ colonies/mg de nodule chez *A. afraspera* (~0,5% de survie). Il est intéressant de noter que USDA110 a formé $1,13 \times 10^8$ colonies/mg de nodule dans *A. afraspera* (~1% de survie), ce qui indique que, malgré l'absence d'agrandissement cellulaire et d'endoréplication, les bactéroïdes USDA110 perdent leur viabilité et subissent une véritable différenciation terminale dans *A. afraspera*. Ainsi, dans la plante *A. afraspera* productrice de NCR, les bactéroïdes USDA110 présentent une déconnexion des quatre caractéristiques canoniques de la TBD (c'est-à-dire la taille des cellules, le niveau de ploïdie, la perméabilité de la membrane et la viabilité cellulaire).

iv. Remarques conclusives

La résistance NCR est à multiples facettes et est connue comme étant nécessaire à la symbiose. Comme montré ici, les rhizobia doivent se défendre pour pouvoir établir une infection chronique dans les cellules symbiotiques des nodules producteurs de NCR. D'autre part, le profil des peptides NCR produits dans les cellules des nodules est également déterminant pour l'issue de la symbiose. Certains mutants de *M. truncatula* dans des gènes NCR individuels ou des accessions de *M. truncatula*, exprimant des allèles NCR spécifiques, présentent une incompatibilité avec les souches de *S. meliloti* (81-84). Il faut donc établir un équilibre délicat dans les cellules nodulaires symbiotiques entre la batterie de NCR sécrétée et les contre-mesures bactériennes multifactorielles correspondantes. Les perturbations chez l'hôte ou chez l'endosymbiote, comme celles décrites ici, qui affectent cet équilibre entraînent une rupture de la symbiose.

La symbiose mal adaptée entre USDA110 et *A. afraspera* constitue un modèle intéressant pour étudier les effets des NCRs sur la différenciation terminale des bactéroïdes. En effet, celle-ci a lieu dans l'hôte *A. afraspera* producteur de NCR, car la viabilité bactérienne est altérée chez les bactéroïdes, tandis que les changements morphologiques et le passage du

cycle cellulaire à l'endoréplication ne sont pas observés. Nous avons également montré, en combinant la protéomique et la transcriptomique, que le programme symbiotique bactérien est exprimé dans les nodules d'*A. afraspera* d'une manière similaire à celle de *G. max*, bien que des modèles spécifiques à l'hôte aient également été identifiés. Cependant, la bactérie semble être en conditions de stress chez *A. afraspera*, possiblement à cause de la production de peptides de type NCR dans cette plante. L'intégration d'ensembles de données provenant de différentes bactéries en symbiose avec un seul hôte, comme ORS285 et USDA110 en symbiose avec *A. afraspera*, a mis en évidence les différences dans les réponses au stress activées chez *A. afraspera* et a confirmé que la symbiose est fonctionnelle mais suboptimale dans cette interaction. Les données moléculaires présentées ici fournissent un ensemble de fonctions candidates qui pourraient être analysées pour leur implication dans l'adaptation à un nouvel hôte et au processus TBD.

OUTLINE

Introduction.....	1
A. Bacteria: invisible to eyesight but with great impact on the living	3
i. The bacterial kingdom.....	3
ii. How can bacteria adapt to new environments ... or to new threats?	5
iii. Bacteria and the management of biological stresses	9
A coating as a shield and camouflage – Surface Polysaccharides of Gram-negative bacteria.	9
The Envelope or Extra-cytoplasmic Stress Response (ESR) - A bacterial Swiss army knife.....	11
iv. Bacteria interacting with the living	12
Bacteria interaction with “higher” organisms and other bacteria	13
Plant-microbe interactions.....	16
B. The mutualistic symbiosis between rhizobia and legume plants.....	19
i. A close interaction between two partners.....	19
What are rhizobia?	19
... and what are legumes?.....	20
General functioning and stakes around this biological process.....	22
ii. The initiation of the symbiosis – A molecular dialog between plant and bacteria.....	25
The plant request – Flavonoids as a “message in a bottle”	25
The bacteria answer – Nod Factors as the trigger.....	26
The notion of host spectrum	30
Eliciting nodule formation – Nod-dependent and Nod-independent symbioses	30
iii. The initiation of the symbiosis – An infection process leading to endosymbiosis	32
The different stages of the infection – Shepherd’s crook model	32
Infection processes in Aeschynomene species.....	34
How is the infection regulated?.....	36
iv. The initiation of the symbiosis – Nodule formation and structure.....	37
Indeterminate nodules of the Medicago genus.....	37
Comparing indeterminate and determinate symbiotic organs.....	39
v. The fixing stage of the process and beyond	40
Nitrogenase – The enzymatic complex responsible for nitrogen fixation	40
Nitrogenase sensitivity to dioxygen – How is it overcome?.....	42
A nutritional symbiosis is a matter of metabolite exchanges.....	44

The lasts stages of the interaction	46
C. The Bacterial differentiation within nodules	48
i. Bacteroid features are modified to reach the TBD	48
ii. NCR peptides as orchestrators of this symbiotic process	51
NCR peptides in IRLC	51
NCR-like peptides in Dalbergioids	58
iii. How TBD impacts symbiosis – The notion of Symbiotic efficiency	59
iv. Bacterial response to differentiation – which functions to survive the NCR burst	60
Envelope functions	60
Transporting NCRs to protect the envelope – The role of BacA	63
Some bacteria totally resist the NCR burst	68
D. Objectives of the thesis project	68
<i>Thesis project studies</i>	71
A. <i>Bradyrhizobium diazoefficiens</i> USDA110 nodulation of <i>Aeschynomene afraspera</i> is associated with atypical terminal bacteroid differentiation and suboptimal symbiotic efficiency.	73
i. Foreword	73
ii. Abstract	75
iii. Importance	75
iv. Introduction	76
v. Results	77
<i>B. diazoefficiens</i> USDA110 is poorly matched with <i>A. afraspera</i> for nitrogen fixation	77
Overview of the USDA110 bacteroid proteomes and transcriptomes	79
Symbiotic functions common to both types of USDA110 bacteroids	82
Host-specific functions	86
USDA110 transcriptomics data in the perspective of previously described regulons and stimulons	87
Expression pattern of orthologous genes between ORS285 and USDA110 in <i>A. afraspera</i> nodules	88
<i>B. diazoefficiens</i> USDA110 bacteroids undergo bona fide TBD in <i>A. afraspera</i> nodules despite very weak morphological and ploidy modifications	90
vi. Discussion	92
<i>A. afraspera</i> triggers atypical but terminal differentiation of USDA110 bacteroids	92

Terminal differentiation is associated with specific stress response.....	94
Correlation between bacteroid differentiation features and symbiotic efficiency for the plant..	95
vii. Conclusion.....	95
viii. Material and Methods.....	96
Bacterial cultures and bacteroid extraction.....	96
Phylogeny analysis.....	96
Genome annotation and RNA-seq analysis.....	97
Proteomic analysis.....	97
Metabolomic analysis.....	98
Plant biomass and nitrogen fixation analysis.....	98
Analysis of <i>B. diazoefficiens</i> USDA110 regulons and stimulons.....	99
Comparison of orthologous gene expression between <i>B. diazoefficiens</i> USDA110 and <i>Bradyrhizobium</i> sp. ORS285.....	99
Analysis of TBD features.....	99
Western blot analysis.....	100
ix. Acknowledgments.....	100
x. Authors' contribution.....	101
xi. References.....	102
xii. Supplementary figures and tables.....	107
xiii. Additional results and discussion.....	112
Functional analysis of similarly regulated genes in USDA110 bacteroids from <i>A. afraspera</i> and <i>G. max</i>	112
Functional analysis of DEG and DAP between <i>A. afraspera</i> and <i>G. max</i> USDA110 bacteroids..	115
Absence of phenotype in target genes mutants - reliability of omics techniques?	116
Modeling as a way to overview the functioning of a metabolism and regulation networks.....	119
B. <i>Sinorhizobium meliloti</i> functions required for resistance to the antimicrobial NCR peptides and bacteroid differentiation	121
i. Foreword.....	121
ii. Abstract.....	123
iii. Importance.....	123
iv. Introduction.....	124
v. Results.....	127
<i>Sinorhizobium meliloti</i> mutants with enhanced sensitivity to NCR peptides	127
Nodule formation by NCR-sensitive <i>Sinorhizobium meliloti</i> mutants.....	129

Can the increased membrane permeability of the NCR-sensitive <i>Sinorhizobium meliloti</i> mutants be attributed to the NCRs in nodules?	131
Bacteroid differentiation in the NCR-sensitive <i>Sinorhizobium meliloti</i>	132
Defective bacteroid differentiation in <i>yejE</i> and <i>yejF</i> mutants	134
Are the hypertrophied bacteroids functional?	136
Peptide uptake by the YejABEF transporter	137
vi. Discussion	139
Multiple functions of <i>Sinorhizobium meliloti</i> contribute to NCR resistance and are required for bacteroid formation and persistence	139
Response to NCR-induced stress regulated by the alternative sigma factor RpoH1	141
The lipopolysaccharide barrier against NCR membrane damage	142
The YejABEF peptide transporter provides resistance to membrane-damaging peptides and sensitizes bacteria to AMPs with intracellular targets	144
vii. Conclusions	145
viii. Materials and Methods	146
Bacterial strains, plant growth and nodulation assays and analysis	146
In vitro sensitivity and peptide uptake assays	147
TEM	147
ix. Acknowledgements	147
x. Author contributions	148
xi. References	149
xii. Supplementary figures	156
xiii. Additional results and discussion	166
Roles of <i>yeyE</i> and <i>yeyF</i> genes in NCR resistance <i>in planta</i>	166
How do BacA and YejABEF importers carry out their activities?	169
Specificity of NCR action	170
General discussion and conclusion	173
i. Cell envelope integrity – a key feature for a successful symbiosis	175
ii. Roles of NCR peptides and evolution of TBD	177
iii. AMPs control biotic interactions	179
iv. The symbiotic genes quest	180
v. Assessing symbiotic efficiency	182
vi. Increasing symbiotic efficiency	183
vii. Conclusive remarks	185

References.....	188
Annexes.....	213
ANNEX A: Generation and phenotyping of USDA110 mutants.....	215
ANNEX B: Tn-Seq, a Transposon Insertion Sequencing technique	217
ANNEX C: Metabolic modelling of ORS285	219

FIGURE LIST

Figure 1 Description of the morphology and the cell boundaries composition of known bacteria.....	2
Figure 2 Tree of life – Phylogeny of fully sequenced organisms.....	4
Figure 3 Schematic representation of horizontal gene transfer (HGT) mechanisms.....	7
Figure 4 Example of an extra-cytoplasmic/envelope stress response (ESR) signaling pathway.....	12
Figure 5 Example of different metazoan-bacteria symbioses.....	14
Figure 6 Illustration of mycorrhizal interactions.....	17
Figure 7 Phylogeny and genome organization of rhizobia.....	18
Figure 8 Legume phylogeny.....	21
Figure 9 Fluxes in the global nitrogen cycle on land and in the ocean.....	23
Figure 10 Molecules involved in the molecular dialog.....	26
Figure 11 Molecular dialog between legume and rhizobia.....	28
Figure 12 Infection of legume roots by rhizobia.....	33
Figure 13 Infection of <i>Aeschynomene</i> species roots by rhizobia.....	35
Figure 14 Comparing determinate and indeterminate nodules.....	38
Figure 15 Genomic structural and functional organization of the nitrogenase.....	41
Figure 16 Regulation of bacterial symbiotic gene expression by O ₂ tension.....	43
Figure 17 Bacteroid metabolism and exchanges with the plant.....	45
Figure 18 Terminal bacteroid differentiation features.....	49
Figure 19 Cell cycle regulation in <i>Sinorhizobium meliloti</i>	50
Figure 20 NCR peptides and their TBD-inducing activity.....	52
Figure 21 Intracellular targets of NCR peptides.....	53
Figure 22 Effects of the mutation of the signal peptidase DNF1 on bacteroid differentiation.....	54
Figure 23 Expression pattern of NCR genes.....	57
Figure 24 Comparison of <i>Aeschynomene</i> and <i>Medicago</i> NCR motifs with defensin motifs.....	58
Figure 25 Envelope functions of bacteroids potentially involved in NCRs.....	61
Figure 26 <i>Sinorhizobium meliloti</i> 1021 <i>bacA</i> mutant in planta phenotype.....	64
Figure 27 <i>Sinorhizobium meliloti</i> 1021 <i>bacA</i> mutant phenotype within <i>M. truncatula</i> <i>dnf1</i> nodules.....	66
Figure 28 The non-adapted symbiotic couple formed by <i>B. diazoefficiens</i> USDA110 and the NCR-producing plant <i>A. afraspera</i> displays suboptimal nitrogen fixation and nodule metabolic dysfunction.....	78
Figure 29 Experimental setup and general description of the transcriptomics and proteomics dataset.....	80
Figure 30 Symbiosis and host-specific functions that display congruency between transcriptomics and proteomics.....	84
Figure 31 Expression pattern of <i>B. diazoefficiens</i> USDA110 and <i>Bradyrhizobium</i> sp. ORS285 orthologous genes <i>in planta</i> and in culture.....	88
Figure 32 <i>B. diazoefficiens</i> USDA110 displays atypical bacteroid differentiation features in <i>A. afraspera</i> nodules.....	91

Figure 33 Functional studies of selected genes displaying no or little symbiotic phenotypes.	113
Figure 34 Functional studies of genes involved in phospholipid synthesis.	114
Figure 35 Functional studies of genes involved in Calvin Benson Bassham pathway.	116
Figure 36 Sensitivity profile of <i>Sinorhizobium meliloti</i> strains to a panel of NCR peptides.	128
Figure 37 Symbiotic phenotype of <i>Sinorhizobium meliloti</i> 1021 mutants during symbiosis with <i>Medicago truncatula</i> wild type and <i>dnf1</i> mutant.	130
Figure 38 DNA content in nodule bacteria in <i>Medicago truncatula</i>	133
Figure 39 Bacteroid morphology of <i>yejE</i> and <i>yejF</i> mutants in <i>Medicago sativa</i> nodules.	135
Figure 40 Nitrogenase expression in the <i>yejE</i> and <i>yejF</i> mutant bacteroids in <i>Medicago truncatula</i> nodules.	136
Figure 41 Uptake of the Bac7 and NCR247 peptide mediated by YejABEF.	138
Figure 42 Symbiotic phenotype of <i>Sinorhizobium meliloti</i> 1021 mutants during symbiosis with <i>Medicago truncatula dnf</i> mutants.	167
Figure 43 Conservation of <i>yejABEF</i> genes synteny.	169
Figure 44 Suggestion of lipid raft (LR) organization.	176
Figure 45 Schematic representation of bottleneck effects.	180

INTRODUCTION

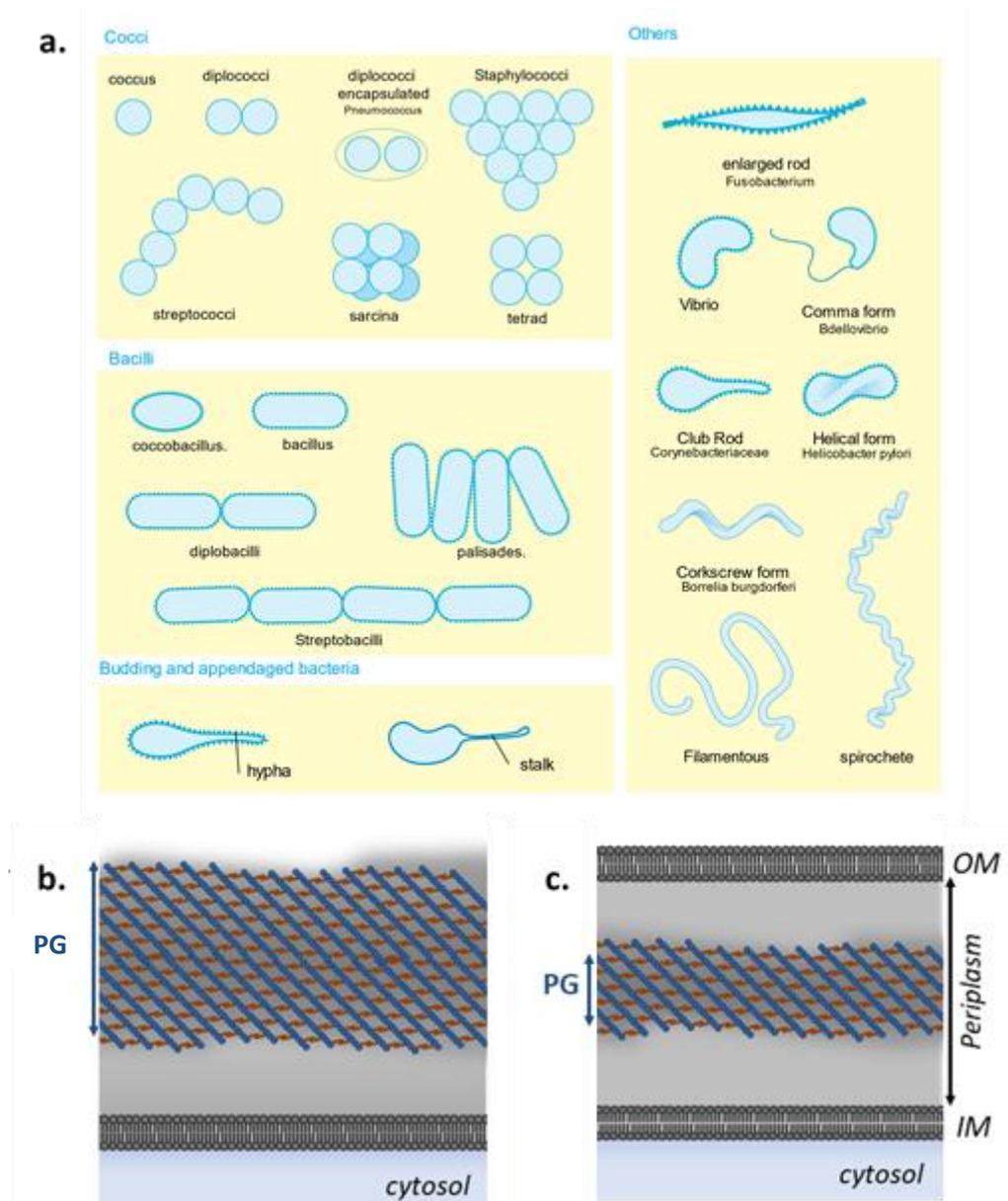


Figure 1 | Description of the morphology and the cell boundaries composition of known bacteria.

a | Morphology of bacteria is diverse ranging from cocci and bacilli, clustered or not to more exotic shapes such as the helical or corkscrew morphologies. *This figure was extracted from https://en.wikipedia.org/wiki/Bacterial_cellular_morphologies.* **b & c** | Structure of bacterial cell boundaries of Gram-positive bacteria (**b**) also called cell wall, or Gram-negative bacteria (**c**) called envelope. IM: inner membrane; OM: outer membrane; PG: peptidoglycan.

B. Bacteria: invisible to eyesight but with great impact on the living

All lifeforms are shaped by their environment and the organisms they interact with. Bacteria, though of microscopic size, often shape their environment and the structure of populations present in the same environment. Indeed, whether in human guts, the soil or plant rhizosphere, bacteria have a huge impact on the living.

i. The bacterial kingdom

Our knowledge of bacterial diversity is far from exhaustive nowadays, as most bacteria have not been characterized yet. Only a small portion of known bacteria are culturable in laboratories (Amann et al., 1995; Hall, 2007). Therefore, some species are only known through sequencing in metagenomic experiments. There are typically 40 million bacteria in a gram of soil and a million in a milliliter of freshwater. It was estimated that approximately 5×10^{30} bacteria are found on Earth (Whitman et al., 1998). Bacteria thus form the second most important world biomass after plants (Bar-On et al., 2018). Though bacteria are unicellular organisms, several morphological and structural features can be used to distinguish them (**Fig. 1a**). Their morphologies can be separated into three major types: cocci (round shaped cells), bacilli (rod-shaped cells), budding and appendaged cells (harboring stalks or hypha). Other types of morphology are found which do not belong to the previous classes such as filamentous cells or corkscrew forms. These morphological criteria as well as the metabolic abilities were first used to classify the bacterial kingdom.

Structural features of bacterial envelope were also used to establish an informal classification (**Fig. 1b&c**). Indeed, the Gram stain procedures allows the distinction of Gram-positive and Gram-negative groups, with a few organisms being Gram-variable (Salton & Kim, 1996). Gram-positive and Gram-negative bacteria display drastic differences in the organization of the structures outside the plasma membrane. Gram-positive bacteria usually harbor a thick (about 20 to 80 nm) layer of peptidoglycan (also called murein) which forms a cell wall (**Fig. 1b**). In Gram-negative bacteria, the peptidoglycan is much thinner (about 5 to 10 nm thick). Moreover, in these bacteria an outer membrane (OM) surrounds the peptidoglycan layer, which itself wraps the inner membrane (IM), resulting in the presence of a cell envelope (**Fig. 1c**). Sometimes, this membrane is anchored with hydrophobic interactions to lipoproteins

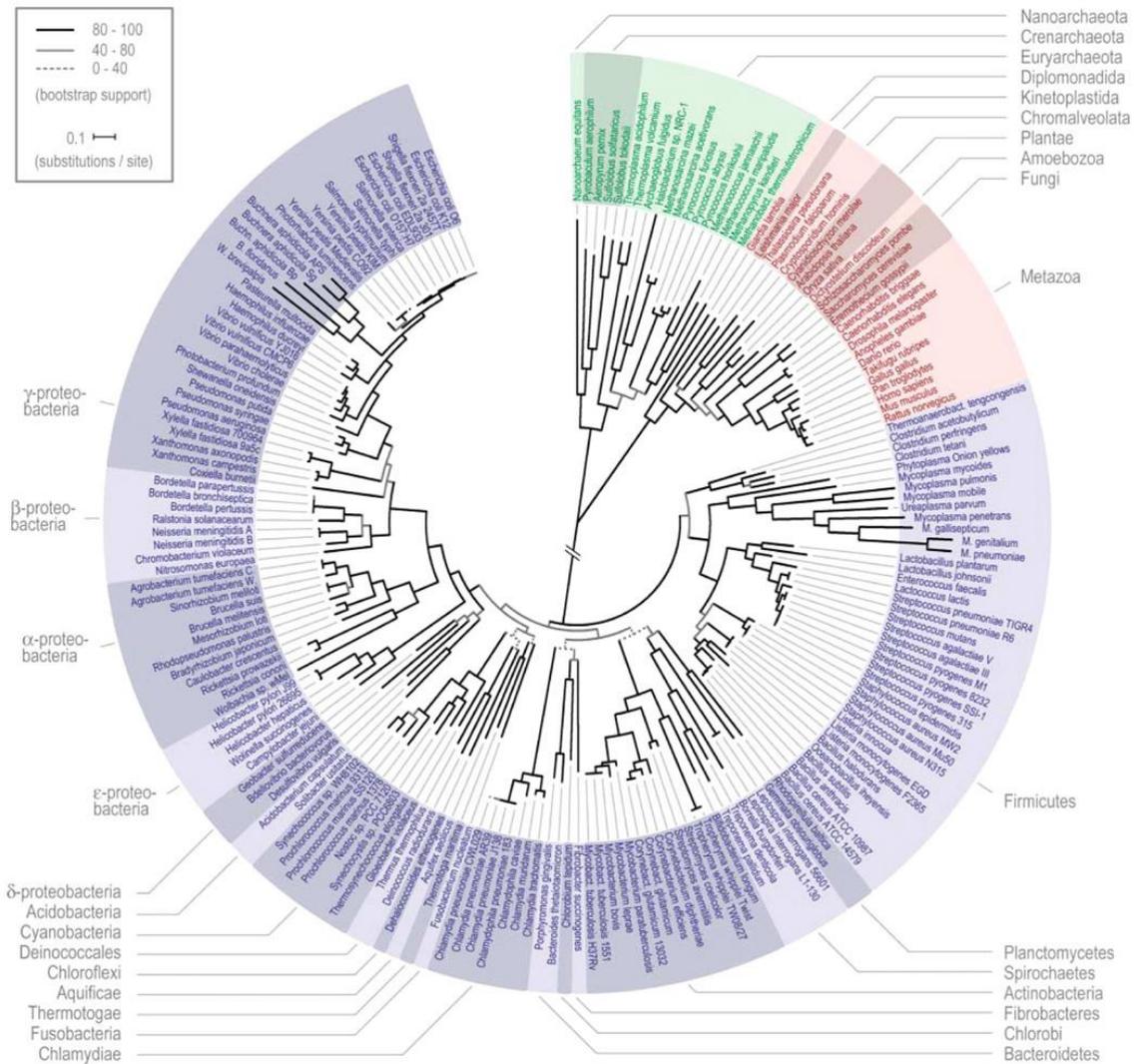


Figure 2 | Tree of life – Phylogeny of fully sequenced organisms. This phylogeny is based on the sequence of 31 ubiquitous proteins present in 191 different species. All the species here have full genome sequence available. Colors segments the tree according to the different kingdoms: green – Archaea; red – Eukaryotes; blue – Bacteria. Color shadings is used to distinguish different well-known subdivisions. Extracted from Cicarelli et al. 2006.

(Braun's lipoprotein), which links the OM with the peptidoglycan (Salton & Kim, 1996). Still in Gram-negative bacteria, the lipopolysaccharides (LPS) are anchored in the outer leaflet of the outer membrane.

In 1987, Eubacteria were divided into 11 *phyla* based on 16S ribosomal ribonucleic acid (rRNA) sequences (Woese, 1987). These *phyla* are composed of purple bacteria, Gram-positive bacteria, Cyanobacteria and chloroplasts, Spirochaetes, green sulfur bacteria, Bacteroides and Flavobacteria, Planctomyces, Chlamydiae, radioresistant Micrococci, green non-sulfur bacteria and Thermotogae (**Fig. 2**). Later, purple bacteria were renamed Proteobacteria, the Bacteroides and Flavobacteria *phyla* were fused and renamed Bacteroidetes and Planctomyces is now called

Planctomycetes. Also, the radioresistant Micrococci are now referred to as *Deinococcus* and green non-sulfur bacteria as Chloroflexi. On **figure 2**, Actinobacteria and Firmicutes are Gram-positive bacteria (Ciccarelli et al., 2006).

The proteobacteria is a phylum containing Gram-negative bacteria. It hosts diverse pathogenic bacterial genera such as *Escherichia*, *Salmonella*, *Vibrio*, *Helicobacter*, *Yersinia*, *Legionellales*, and many others (Madigan & Martinko, 2006). This division was separated into nine sub-divisions, also based on 16S rRNA. It is composed of Alpha-, Beta-, Gamma-, Delta-, Epsilon- and Zetaproteobacteria, as well as Hydrogenophilalia, Acidithiobacillia and Oligoflexia. Some Alphaproteobacteria are known to be able to grow with low levels of nutrients and can have special morphology features such as stalks and buds. Other genera in this class include agriculturally important bacteria such as plant pathogens and symbionts (*Agrobacterium*, *Rhizobium* or *Sinorhizobium* for example), as well as animal pathogens (*Brucella abortus*). The Betaproteobacteria display a high diversity in the metabolic functions represented. They notably contain generalist heterotrophs and Burkholderiales, which comprise animal and plant pathogens or symbionts. The Gammaproteobacteria class comprises numerous plant pathogens such as *Pseudomonas*, *Xanthomonas*, *Erwinia*, *Dickeya*, *Xylella* and *Pectobacterium*. In the rest of the manuscript, except if otherwise stated, bacteria will essentially be referring to Gram-negative bacteria.

ii. How can bacteria adapt to new environments ... or to new threats?

Bacteria are highly adaptive organisms that can colonize a wide range of environments. Beside our tempered-climate environment, they can either be found in niches with high physical constraints such as cold or desiccation, or with chemical constraints such as alkaline, acidic, or saline environments. Some bacteria can even be highly resistant to radiations such as UV or ionizing radiations. This high versatility observed in the bacterial kingdom is rarely transposable in animal or plant kingdoms. Indeed, whereas some plant species are adapted to high or low temperatures, saline soil, and some fish species to high pressures they rarely survive as extreme conditions as bacteria do.

This ease with which bacteria can colonize new niches can be explained by several features. For example, their rather short duplication time leads to an accumulation of

spontaneous mutations, as they can be transmitted rather quickly to the offspring compared to animals or plants. Another bacterial feature that contributes to this high mutability is the ploidy level of bacterial genomes. Indeed, bacteria are haploid organisms meaning they only have one copy of their chromosomes compared to plants or animals. This property enhances their adaptability as bacteria cannot benefit from DNA repair mechanisms such as homologous repair (HR) that involve a second copy of the chromosome to be used as a template for correcting mutations. These features lead to an increased ease to fix mutations into bacterial populations. Speaking of quick accumulation of mutations is very subjective as mutation rate vary from one organism to the other and depending on the considered sequences. As an example, a study of long-term *Escherichia coli* evolution showed mutation rates can vary from 10^{-10} to 10^{-7} mutations per generation (Sniegowski et al., 1997). This is not the safest way to adapt to a new environment, from an anthropomorphic point of view. As mutations occur randomly, they thus generate either beneficial, neutral, or deleterious effects. The likelihood of a beneficial mutation can be low, even under selection pressure. Nonetheless, some studies have shown that in this way bacteria can become quite quickly resistant to a high concentration of antibiotics. Indeed, Baym and colleagues have shown in 2016 that within 264 hours, *E. coli* K12 variants were able to adapt and grow in media containing trimethoprim at sequential increasing concentrations of 3, 30, 300 and 3000 times the Minimal Inhibitory Concentration (MIC) of the Wild-Type (WT) strain. In the end, a total of 16 independent clones could reach and grow on the most concentrated part of the plate.

However, spontaneous mutation is not the most powerful and quickest way, for bacteria to adapt to different environments. The feature of bacteria adaptability which is mainly responsible for adaptation is the horizontal gene transfer (HGT), which also allows acquiring new complex metabolic abilities. Bacteria can exchange different sizes of DNA pieces through three main processes: conjugation, transformation, and phage transduction (**Fig. 3**).

Transduction by bacteriophage is an erroneous functioning of the viral infection process (**Fig. 3**) (Braga et al., 2018). Bacteriophages are viruses targeting bacteria. While they usually use their host to replicate before lysing it, sometimes, they can integrate bacterial genomic fragments. Their size is limited by the capsid (the phage envelope) capacity. By

infecting a new cell, the phage can transfer its DNA to a receiver cell, which will integrate it in its genome using the phage machinery.

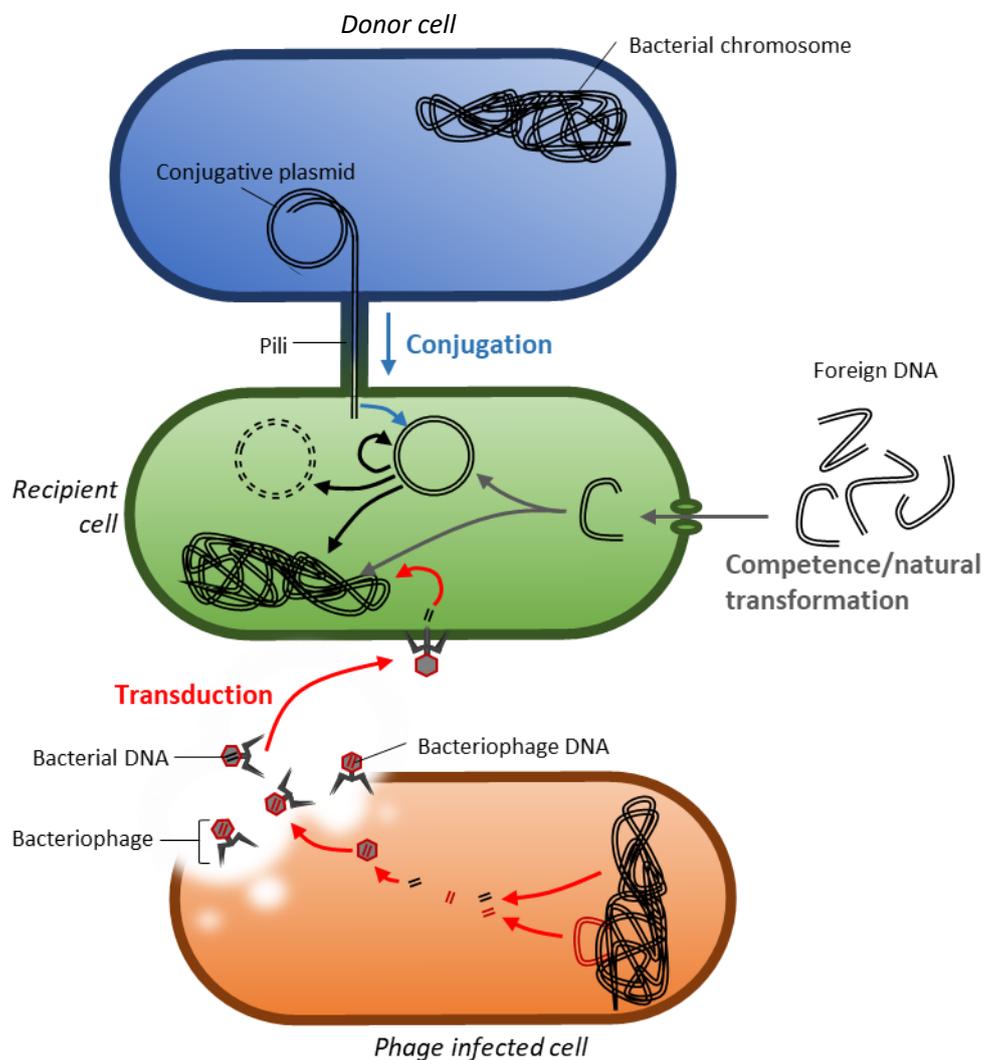


Figure 3 | Schematic representation of horizontal gene transfer (HGT) mechanisms. Bacteria have several ways of integrating exogenous DNA: conjugation, transformation, and transduction by bacteriophages. The green central cell represents a recipient cell for all three mechanisms. Blue arrows represent the conjugation steps, grey is for transformation and red for transduction. Black arrows represent the fate of a newly integrated plasmid in a bacterial cell. Inspired from von Wintersdorff et al, 2016.

Transformation is the process by which bacteria can acquire naked DNA present in its environment when it is in a permeable state (**Fig. 3**). In that case, DNA has a chance to penetrate through the pores formed in the envelope of the bacteria. These pores made thanks to complex machinery allows the exogenous DNA to reach the cytosol (Johnston et al., 2014). This machinery comprises a retractile pilus used to bind and capture the exogenous DNA, as well as translocases to allow the passage of the DNA through the membranes (Johnston et al.,

2014). In the cytosol the exogenous DNA can either be integrated into the genome, replicated, degraded, or lost during generations. Several conditions are required to perform a transformation. An exogenous DNA must be present near the recipient cell, which must be in a competent state. The internalized DNA has to then be stabilized either by circularization or integration in the recipient cell genome (Thomas & Nielsen, 2005). Competence can be induced in specific conditions depending on the sensing of the nutrient status, the growth phase of the culture or a physiological stress as an example (Johnston et al., 2014). The transformation process was described to occur not only in laboratories but in nature as well such as for the acquisition of antibiotic resistance genes (Pletz et al., 2006). However, these natural occurrences are rarely observed, and it is unknown for most bacteria if they happened to be naturally transformed and what would be the required signals or conditions to induce competence.

Through conjugation, DNA, usually plasmids, can be transferred from one bacterium to another (**Fig. 3**). This transfer requires a physical contact between the donor cell and the recipient cell via cell surface pili or adhesins (Von Wintersdorff et al., 2016). To perform the conjugation, the donor strain must have functional *mob* and *tra* operons and the plasmid must carry a transfer origin sequence. This machinery is usually encoded by genes on plasmids or integrative conjugative elements in the chromosome (Smillie et al., 2010; Wozniak & Waldor, 2010). Antibiotic resistance genes are often associated with conjugative elements (plasmids or transposons) (Von Wintersdorff et al., 2016). Conjugation might certainly not be the only HGT method to acquire antibiotic resistance genes. However, it is considered as the most likely mechanism out of the three. Indeed, conjugation combines several advantages compared to the other mechanisms (Von Wintersdorff et al., 2016). First, the transferred genetic material is protected from the surrounding environment, as it never leaves the cytoplasm of a cell. Then, the mean to enter the recipient cell is more efficient than in transformation. Moreover, conjugation is a bacterium-borne process dedicated to the transfer of bacterial genes, whereas the transfer of bacterial DNA by transduction is due to an abnormal bacteriophage replication (Norman et al., 2009). Transfer of plasmids and conjugative transposons, sometimes between distant bacteria, has been described several times (Musovic et al., 2006; Shoemaker et al., 2001; Tamminen et al., 2012). Conjugation is also at the origin of the spreading of various antibiotic resistances in pathogen populations (Huddleston, 2014).

Compared to natural transformation and transduction, conjugation can lead to the integration of large sets of genes into the receiver genome. For instance, it has been shown that conjugation is responsible for the dissemination of the virulence plasmid pTi of the *Agrobacterium tumefaciens* plant pathogen (Lang & Faure, 2014). Thus, acquiring new essential function, or allowing a better adaptation to carbon or nitrogen sources available in its environment. Besides colonizing a new environment, these changes can help the bacterium to be more competitive in an environment or even to build new interactions with other organisms.

iii. Bacteria and the management of biological stresses

To face the large diversity of stresses that can be encountered in their environment, bacteria evolved a wide range of adaptive strategies, including metabolic adaptations, as well as structural modifications. Rather than exhaustive, this section will present a selection of bacterial functions often recruited for stress resilience that will be discussed later in this manuscript.

A coating as a shield and camouflage – Surface Polysaccharides of Gram-negative bacteria.

The envelope of Gram-negative bacteria can be arranged in a complex mixture of layers, which usually all play part in host-bacteria interaction. Whether in pathogenic or symbiotic interactions, all the components described here either hide bacteria from their host surveillance systems or help the bacteria to survive host defenses.

A characteristic feature of Gram-negative bacteria is the presence of LPS covering the OM. These components are anchored to the outer leaflet of the OM through a hydrophobic moiety and a large hydrophilic part faces the extracellular medium. The LPS possesses an array of powerful biologic activities and play an important role in the pathogenesis of many Gram-negative bacteria. LPS can be composed of three regions: the lipid A which anchors the LPS in the outer membrane, the LPS core covalently bound to the lipid A and the O-antigen which are polysaccharide chains linked to the core (Steimle et al., 2016). The core is composed of various saccharides (heptose, glucose, galactose) and 2-keto-3-deoxyoctonic acid (KDO), ethanolamine, and *N*-acetylglucosamine. The individual monosaccharide components of O-

antigens confer serologic specificity to the LPS. This capsular polysaccharide is usually assembled at the IM before being transported to the OM (Salton & Kim, 1996). As surface polysaccharides, LPS are one of the first protections bacteria harbor. Therefore, they are usually highly immunogenic during pathogenic interactions (Steimle et al., 2016). Structural modifications can be performed to alter this immunogenicity though, such as in *Pseudomonas aeruginosa* (Maldonado et al., 2016) an opportunistic pathogen. For instance, the loss of the O-antigen can favor the infection by lowering the immune response.

Bacteria also secrete molecules known as exopolysaccharides (EPS) to protect themselves from the environment, avoid recognition by plant defenses or create biofilms, a bacterial community. EPS are homo- or heteromeric polysaccharides, secreted by numerous bacteria, forming an envelope around the cell. Their structure is affected by two factors. First, the elementary compounds which are diverse. The most common ones are the saccharides D-glucose, D-galactose and D-mannose, but other types of modifications can also be found, such as succinate, pyruvate and acetate (Whitfield et al., 2015). The second factor that affects the structure of EPS is the level of polymerization, which affects the reticulation of the mesh. EPS have been shown to be involved in protecting the bacteria from host defense responses in pathogenic interactions. Indeed, they can mask the LPS coating, decreasing the cell immunogenicity, or the structure of EPS can be modified to avoid recognition by host surveillance systems (Whitfield et al., 2015). Also, it has been shown that they can protect the bacterial population from host antimicrobial peptides (Karygianni et al., 2020).

KPS (K-antigen polysaccharides) or K-antigen is another type of surface polysaccharide, also known as capsular polysaccharides (CPS). They can be found in numerous bacterial species and have various structures. Indeed, KPS are often secreted as a mixture of K-antigens which are mostly composed of acidic saccharides (Cescutti, 2010). They are usually involved in interactions with a host. In pathogeny, they are often considered as virulence factors. Indeed, their presence helps to protect the bacteria against the host defenses, whether by avoiding the contact with phagocyte in animals due to their negative charge, or by averting interaction with antimicrobial peptides (AMP) (Cress et al., 2014).

Cyclic Betaglacans (CG) are another family of extracellular polysaccharides found almost exclusively in the *Rhizobiaceae*, which belong to the Alphaproteobacteria class. CGs are glucose

polymers, with glycosidic linkages that vary depending on the species. It has been shown they can be important for a successful infection of host plants or for the virulence of plant pathogens (Breedveld & Miller, 1994).

The Envelope or Extra-cytoplasmic Stress Response (ESR) - A bacterial Swiss army knife

As all organisms, bacteria can perceive changes in their environment and react by modifying the expression of genes or their metabolism. Pathogenic or symbiotic bacteria need to be able to perceive their host and to perceive defense reactions before they are eliminated due to the host defenses. As already mentioned, the envelope is the first structure of contact between a signal or a stress and the bacterium. Therefore, several molecular systems have been described to perceive and induce changes in bacterial physiology. They are known as ESR, short for extra cytoplasmic/envelope stress response. Usually, these systems involve a complex sensor machinery, which will lead to the release or the activation of either a transcription factor or sigma factors (Hews et al., 2019). Sigma factors are subunits of the RNA polymerase complex, which are involved in the DNA-binding specificity of the complex. These systems can detect a large array of stresses and environmental changes. For example, they sense shifts in temperature, light, exposure to certain molecules, nutrient availability, and type, as well as being involved in pathogenesis. One ESR system, which is quite polyvalent in terms of the range of stresses detected, is the Cpx signaling cascade in *E. coli* (**Fig. 4**) (Hews et al., 2019; Rowley et al., 2006).

The most studied transcriptional regulator is the sigma factor σ^E or *rpoE*, which is regulated by different ESR systems. It controls the expression of a large array of genes such as proteases, chaperonins, transporters, membrane synthesis/modification enzymes, primary metabolism as well as other sigma factors (Rowley et al., 2006). For example, *Salmonella enterica* serovar Typhimurium σ^E is essential for the protection against reactive oxygen species (ROS) and AMP and possibly through the functions it induces (Rowley et al., 2006). Though not all mechanisms are understood yet, AMP transporters are amongst the genes regulated by σ^E in some species (Delory et al., 2006; Rowley et al., 2006). Other known sigma factors are *rpoD* (σ^D , σ^{70}), *rpoH* (σ^H , σ^{32}) or *rpoN* (σ^{54}). *rpoD* is responsible for the regulation of housekeeping

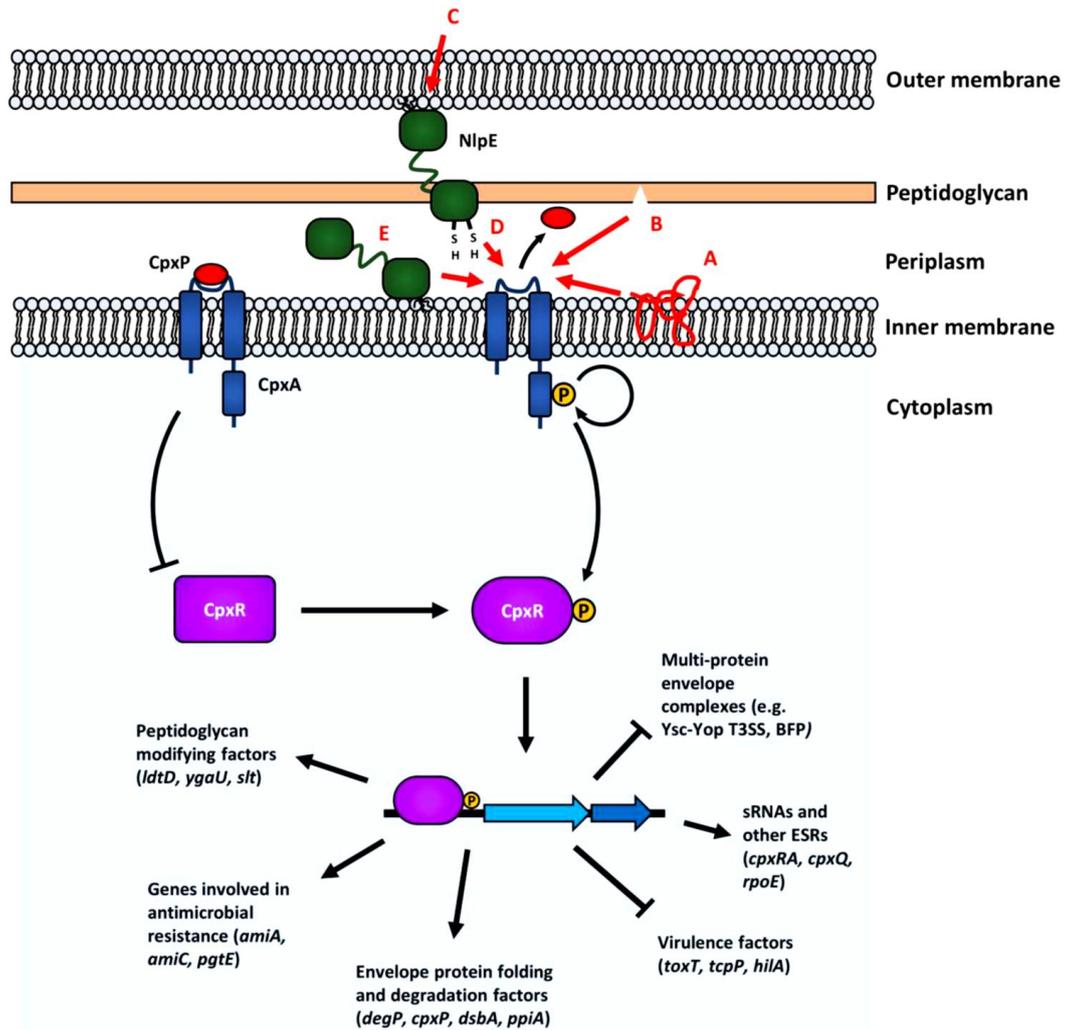


Figure 4 | Example of an extra-cytoplasmic/envelope stress response (ESR) signaling pathway. The Cpx system can sense different envelope stresses. Signals inducing Cpx response are shown with red arrows and associated letters: (A) misfolded proteins; (B) peptidoglycan damage; (C) surface signal; (D) periplasmic redox states; (E) OM lipoproteins trafficking defects. In absence of signal, the CpxA phosphatase is inactive, inhibited by CpxP. However, when a signal is perceived, CpxP is freed from CpxA which phosphatase activity is restored. CpxR is then phosphorylated by CpxA, becoming an active transcription factor. It will activate different target genes. The list presented here is representative but not exhaustive of CpxR activity. Extracted from Hews et al., 2019.

functions, *rpoN* for nitrogen metabolism and *rpoH* which is a heat-shock associated sigma factor though it manages membrane stresses and virulence as well (Rowley et al., 2006).

iv. Bacteria interacting with the living

Herein, I will focus on beneficial interactions. Some of the firsts and ancestral beneficial interactions between bacteria and eukaryotes paved the way towards the evolution of primary and secondary endosymbioses. These processes led to the permanent internalization of the

symbiont, which became obligatory. Intracellular domestication of Alphaproteobacteria formed the primary endosymbiosis and led to the acquisition of energy-producing mitochondria. A second endosymbiosis led to the emergence of the green lineage with the engulfing of a photosynthetic cyanobacterium thus becoming plastids (McFadden, 2014; Roger et al., 2017). These two primary endosymbioses were not the only occurrence of such ultimate symbiosis, which leads to the emergence of a new organism. Indeed, secondary endosymbioses occurred afterwards, generating new cell lineages (Delwiche, 1999). For example, it allowed to transfer the ability to perform photosynthesis into different eukaryote phyla, which resulted in the polyphyletic group of algae. Though these interactions can be seen as the ultimate forms of beneficial symbioses, many described symbiotic interactions are far to reach this state.

Bacteria interaction with "higher" organisms and other bacteria

Bacteria can interact with a wide range of hosts. In fact, they are present in different organs of animal species. For example, the human skin hosts a large diversity of bacteria that help to protect from pathogen aggressions in exchange of nutrients collected from the upper layers of the epidermis (Byrd et al., 2018). Indeed, it has been shown that infection by *Staphylococcus aureus* is avoided thanks to skin microbiota, which degrades the adhesion proteins secreted by the pathogen. Microbiota refers to the whole community of microbes that lives in a specific environment. Another well studied microbial community is the human gut microbiota, which helps digesting some nutrients or (again) to protect from aggressors (Fan & Pedersen, 2020). Generally, in a host-specific microbiota, bacteria reside in a privileged niche in which nutrients are available and where they can multiply more easily in exchange for some help provided to their host. This help is either given in the form of a novel metabolic function that the host cannot perform, or a protective role against pathogenic bacteria.

Thus, bacteria are found in symbiosis with "higher" organisms. A symbiosis is a close, long-term interaction between two organisms. This term englobes parasitism, commensalism, and mutualism. Whereas parasitism defines interactions in which one species takes advantage of the interaction by negatively affecting its partner, in commensalism the partner does not suffer negative effects. However, commensalism appears to be often either parasitic or symbiotic because of our lack of understanding of the concerned interactions. The last type of symbiosis, mutualism, defines associations where both partners benefit from the interaction.

From now on, the term symbiosis will be used to describe mutualistic symbioses. Mutualistic symbioses are quite diverse in terms of species involved and outcomes. Indeed, symbioses are not restricted to bacteria with eukaryotes as multiple types of couples exist. In nature, interactions between eukaryotes (insects with plants such as entomophile pollination or lichens – algae with fungi), prokaryote-prokaryote (syntrophy or complementarity of metabolism) (Morris et al., 2013) or microbes with plants/metazoans are observed.

As for other types of associations, examples of bacteria interacting with eukaryotes are diverse. In the nutritional symbiosis between *Riptortus pedestris*, a Japanese stinkbug pest of pea and soybean crops, and the Betaproteobacterium *Burkholderia insecticola*, the bacterial symbiont is cultivated by its host in specialized gut compartments (**Fig. 5a&b**). The insects are believed to benefit from nutrients provided after digestion of bacteria. This impacts the fertility and development of the stinkbug (Kikuchi et al., 2007). This symbiosis displays another striking feature as it was shown that the symbiont could confer pesticide resistance to the stinkbug (Itoh et al., 2014; Kikuchi et al., 2012). In counterpart, the bacteria benefit from a privileged niche as well as an opportunity to increase its dissemination in the environment after the death of the insect and decay of its carcass.

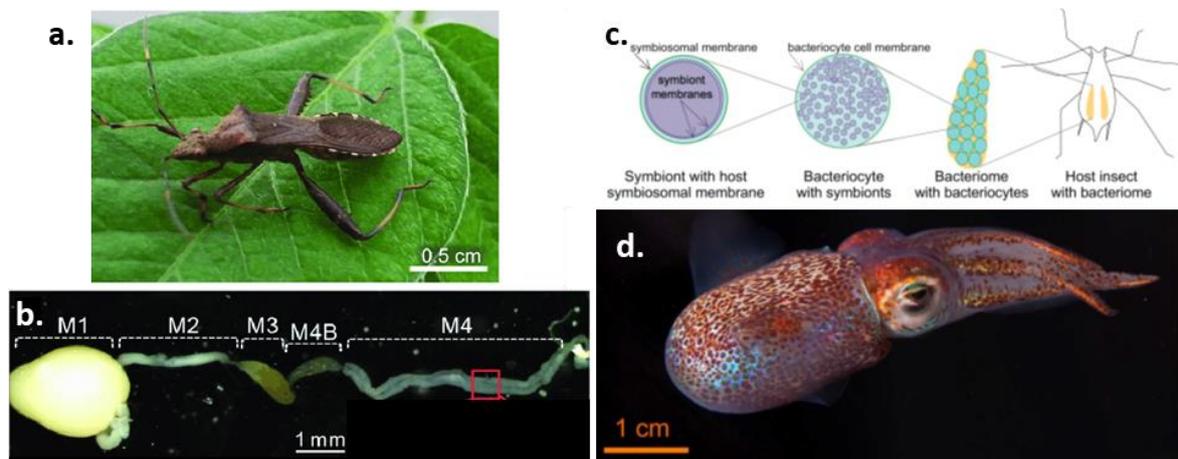


Figure 5 | Example of different metazoan-bacteria symbioses. a & b | The *Riptortus pedestris* – *Burkholderia insecticola* symbiosis. The insect (**a**) hosts thousands of bacteria in its gut (**b**). While M1 to M3 regions are dedicated to digestion, M4B and M4 regions are specialized in hosting symbiotic bacteria. M4 region is composed of crypts where bacteria are cultured by the insect. When bacterial density is high enough, bacteria migrate towards the M4B region where they are digested (Futahashi et al., 2013). Extracted from Kim et al., 2013. **c |** Scheme representing the different levels of encapsulation of *Buchenra aphidicola* symbiotic bacteria in aphids. Symbiotic bacteria are surrounded by host membrane. Multiple bacteria are hosted in host symbiotic cells called bacteriocytes. These bacteriocytes are themselves the component of the symbiotic organ, a tissue called bacteriome. Extracted from Wilson & Duncan 2015. **d |** Hawaiian bobtail squid (*Euprymna scolopes*) in interaction with bioluminescent *Vibrio fischeri*. Low luminescence coming from the ventral part of the squid can be observed by transparency. Extracted from McFall-Ngai 2014.

Another well-studied symbiosis between insects and bacteria is the aphid – *Buchnera aphidicola* nutritional symbiosis. In the aphid, like in many other insect symbioses, the symbionts are present in highly specialized cells, called bacteriocytes (**Fig. 5c**) (Wilson & Duncan, 2015). Bacteria are directly transferred from the mother to her offspring during the gastrula stage of the embryo. This process is called vertical transmission of the symbionts. In comparison, the *Riptortus* – *Burkholderia* symbiosis is transmitted horizontally, as the eggs are free of bacterial symbionts and the insect must acquire its symbiont in the environment at each generation. In the aphid – *Buchnera* symbiosis, the vertical transmission led over long evolutionary times to a very strong reduction of the bacterial genome (Hansen & Moran, 2011). *B. aphidicola* is therefore unculturable in laboratory conditions as it does not survive outside its host.

Besides nutritional symbioses, other types of interactions exist such as the *Euprymna scolopes* – *Vibrio fischeri* (**Fig. 5d**). *E. scolopes* is a bobtail squid, which lives in shallow reef flats of the Hawaiian archipelago. During the day, this squid buries itself in the sand, staying quiescent until dusk, where it emerges and searches for food. One of this squid particularities is the emission of light from a specific organ placed deep in his ventral cavity (McFall-Ngai, 2014). The function of the light emission is to produce an adjustable counter-illumination that masks the animal via silhouette reduction when it is foraging at night under moonlight. The light is produced by a bioluminescent bacterium called *V. fischeri* (Gammaproteobacteria). Every evening, symbionts proliferate in the crypts of the symbiotic organ. There, they form biofilms and will eventually induce their luminescence-producing functions (McFall-Ngai, 2014). A luciferase is responsible for this luminescence, encoded by the *lux* operon. This expression of this operon relies on a community-controlled regulation process called quorum-sensing (QS) (Lupp & Ruby, 2005; Whiteley et al., 2017). This system allows bacteria to tweak their functions based on a population density. N-acyl homoserine lactones (AHL) are the signal molecules used for assessing population density (Whiteley et al., 2017). At dawn, the bacteria are ejected from the symbiotic organ into the ocean, thereby switching off the light organ. This rhythmic acquisition and release of symbionts have been shown to induce the circadian cycles which are important for regulating the day-night life functions of organisms (McFall-Ngai, 2014). In other

marine organisms interacting with *V. fisheri*, the light organ can be used as a *lamparo* (a fishing light attractor) to attract preys (Haygood & Distel, 1993).

Plant-microbe interactions

Plants are also permanently interacting with microorganisms, from the seed stage to fully mature and reproducing plants. Indeed, bacteria are massively present in the soil as stated previously. Thus, plants host organ-specific microbiota which composition can vary during their different development stages. Plants can strongly modify soil microbiota through molecular signals and root exudates (Compant et al., 2019; Müller et al., 2016), but their genotype and architecture also contribute to the selection of their microbiota. Several examples of a biocontrol effect from the plant microbiota over pathogens have been described. Bacterial communities have different tools to do so, such as production of antibiotics and lytic enzymes as examples. Genera such as *Pseudomonas*, *Streptomyces*, *Bacillus*, *Burkholderia* and *Paraburkholderia* have been demonstrated to play part in pathogen suppression (Compant et al., 2019).

Beside these interactions, higher plants are known to be able to carry out two more specific types of symbioses with soil microbes. One of them are nitrogen fixing symbioses. Though several variations of this kind of symbiosis exist in plants, I will focus on the legume-rhizobia symbiosis. This nutritive symbiosis, which emerged about 70 My ago (Strullu-Derrien et al., 2018), will be extensively described in the following chapter. Another one is a symbiosis with fungi, called mycorrhization. This symbiosis concerns 85% of plants and their symbionts are part of the Eumycetes clade of fungi. During this symbiosis, fungal hyphae act as an extension of the root network, increasing the soil volume in which the plant can collect nutrients such as phosphate, nitrogen or potassium. In exchange, the plant gives carbon-based nutrients to the symbionts, primarily sugars and in some cases lipids. Different types of mycorrhiza exist, with different degrees of intimacy in the colonization of the host. Ectomycorrhiza coat some of the roots and hyphae penetrate intercellularly the plant root cortex (**Fig. 6a**). It seems that this type of mycorrhiza appeared 50 My, convergently in different fungi clades whereas on the plant side both convergent evolution or single origin scenarios are still considered today (Strullu-Derrien et al., 2018). However, ectomycorrhiza are restricted to temperate regions and only interact with few ligneous species from gymnosperm and

angiosperm families. Endomycorrhiza or arbuscular mycorrhiza (AM) are the second type of mycorrhiza, which is the most wide-spread and studied mycorrhizal association. Fungi involved in these interactions are from the Glomerales order. AM implies an intracellular colonization of plant cells by arbuscular (tree-like shape) structures generated from hyphae developed in the apoplast (**Fig. 6b**). These ramified structures allow a substantial exchange surface between the plant and the fungus. This symbiosis is much older than the other above described interactions, as it appeared possibly earlier than 400 My ago (Strullu-Derrien et al., 2018), and was concomitant and probably a key step in the colonization of land by plants. Nowadays, AM symbiosis is present from *Marchantiophyta* to angiosperms. Interestingly, some of the signaling mechanisms involved in setting up these different types of symbiotic interactions seem conserved on the plant side and convergent on the microbial side (see chapter B.ii).

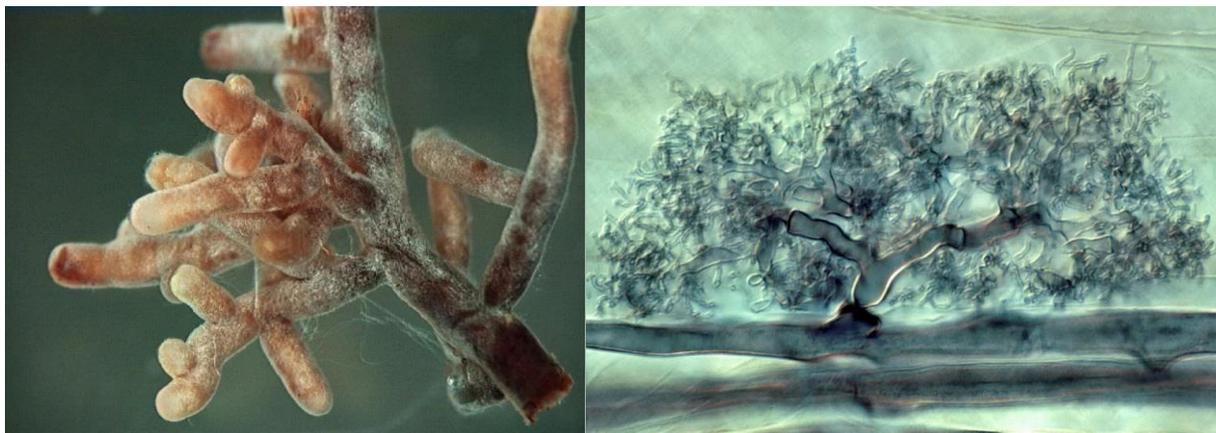


Figure 6 | Illustration of mycorrhizal interactions. a | *Amanita* type ectomycorrhiza. The lateral root architecture is deeply modified (short roots with high level of dichotomy) and the mycelium forms a mantle around the symbiotic roots. Extracted from Nilsson et al, 2005. **b** | Plant root cell filled with an arbuscule from a *Glomus* fungus. The surface of exchange with the host is maximized thanks to the architecture of the interface. The trunk (T) is visible as well as numerous fine dichotomic branched hyphae (arrows). From <http://mycorrhizas.info/vam.html#>

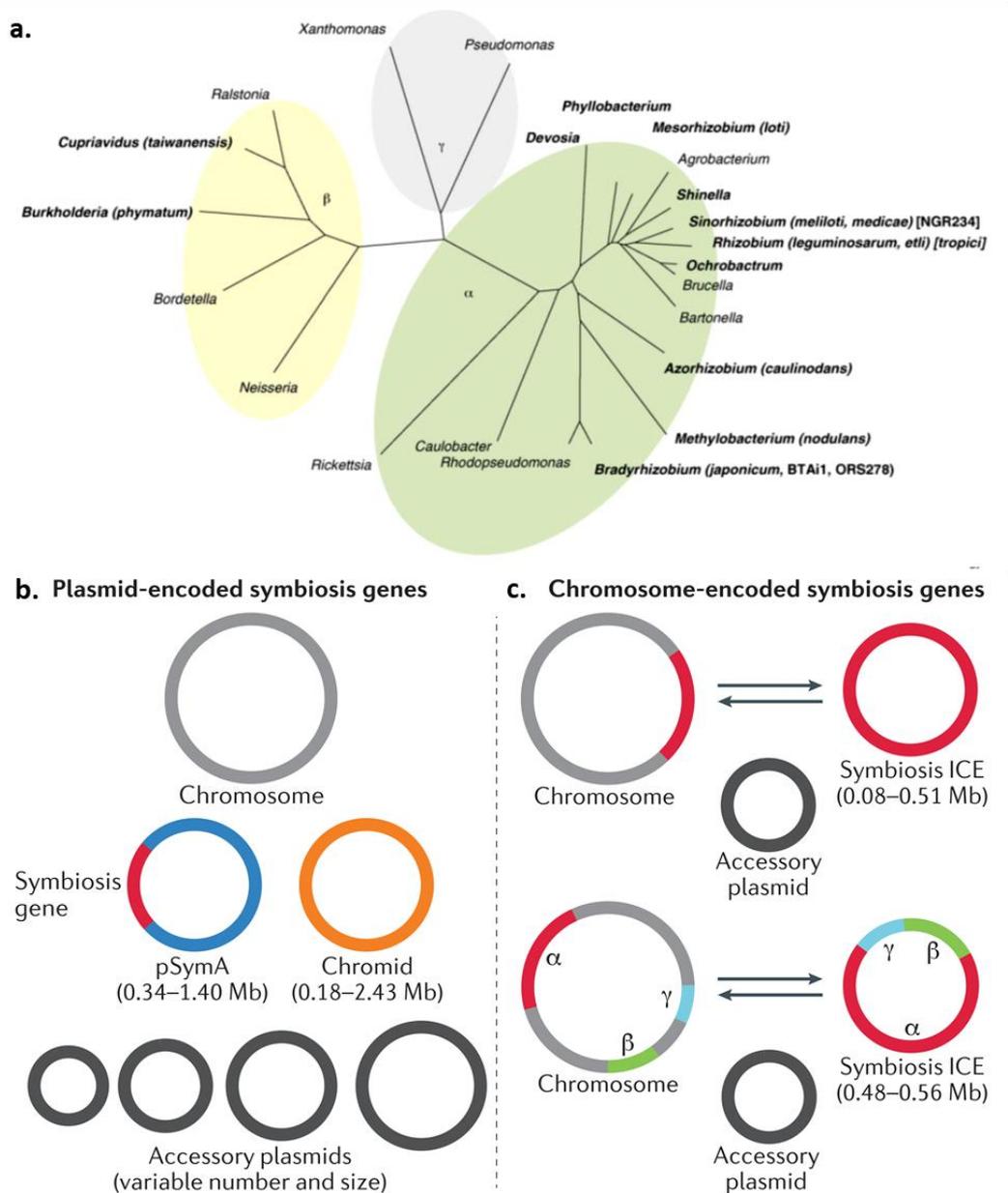


Figure 7 | Phylogeny and genome organization of rhizobia. The polyphyletic clade of rhizobia includes bacteria from distant families. This feature is explained by the high mobility of symbiotic genes. **a** | Phylogenetic tree bacterial groups including different rhizobial genera as well as other bacterial models. The symbiotic rhizobia (bold font) and are found in alpha- and betaproteobacteria. Extracted from Masson-Boivin & al. 2009. **b** | Plasmids baring symbiotic genes can be quite large, up to 1,4 Mb for pSymA in *Sinorhizobium meliloti*. This organization of the genome, including chromosome(s), with additional chromide(s) and/or accessory plasmid(s) is classical among rhizobia. Genera *Rhizobium*, *Sinorhizobium*, *Cupriavidus* or *Burkholderia* display such genome organization. **c** | Other strains have chromosome-encoded symbiotic islands. *Azorhizobium*, *Bradyrhizobium* and *Mesorhizobium* are examples of genera with this type of organization. In some genera, symbiotic genes are localized in ICEs. This specificity allows reshuffling of the symbiotic functions as they can be excised from the chromosome. They will form a plasmid-like replicon, that can afterwards be transferred by conjugation. In the recipient cell, the ICE will be able to integrate the chromosome through aminoacyl-tRNA sequences. Though this behavior was observed in *Azorhizobium* and *Mesorhizobium*, no such feature was noted in *Bradyrhizobium*. **b** & **c** extracted from Poole & al. 2018.

C. The mutualistic symbiosis between rhizobia and legume plants

In this section, I will present the molecular processes that underlie the legume-rhizobium symbiosis for several studied models. As my work was mostly focused on the functions of the bacteria during the interaction, I insisted on depicting the bacterial side.

i. A close interaction between two partners

Before discussing the purpose of this interaction, the two next parts will present the types of organisms that are involved in this specific symbiosis. This interaction involves two partners from well-defined groups. On one side are the legumes or *Fabaceae*, which is the only group that can host these symbionts. On the other side are rhizobia the only bacteria able to induce the formation of a symbiotic organ on the legume plants and to colonize it.

What are rhizobia? ...

The term rhizobia designate a polyphyletic group of bacteria, which can establish a nitrogen-fixing symbiotic interaction with legumes. This group is essentially composed of Alphaproteobacteria but a few Betaproteobacteria also belong to them (**Fig. 7a**) (Moulin et al., 2001). Rhizobia are found in 16 genera, which are phylogenetically distant: *Aminobacter*, *Azorhizobium*, *Bradyrhizobium*, *Devosia*, *Mesorhizobium*, *Methylobacterium*, *Microvirga*, *Neorhizobium*, *Ochrobacterium*, *Pararhizobium*, *Phyllobacterium*, *Rhizobium*, *Shinella* and *Sinorhizobium/Ensifer* for the Alphaproteobacteria plus *Burkholderia* and *Cupriavidus* in the Betaproteobacteria (Masson-Boivin et al., 2009). Today we count more than 150 identified rhizobial species. As a polyphyletic group, the cited genera are not isolated in a unique branch and can comprise non-symbiotic bacteria (Sy et al., 2001). Even plant pathogens are part of this clade, such as *Rhizobium rhizogenes*, formerly known as *Agrobacterium rhizogenes* (Young et al., 2001).

The ability of a bacterium to establish a symbiosis is dependent on the presence of well-described genes in the symbiont genome. These functions will be described later in this manuscript. Interestingly, symbiotic genes are usually clustered in transferable regions of the bacterial genome. For example, in *Sinorhizobium meliloti*, they are distributed over symbiotic mega plasmids or chromides (**Fig. 7b**) (Poole et al., 2018). In other cases, like in *Mesorhizobium*

loti or *Bradyrhizobium diazoefficiens*, they are clustered in specific chromosomal regions called symbiotic islands (**Fig. 7c**) (Kaneko et al., 2000, 2002), which are surrounded by insertion sequences (Iida et al., 2015). In both cases, symbiotic genes can easily be transferred from one strain to another. Indeed, plasmids can be shared through conjugations and symbiotic islands are mobile genetic elements too. Symbiotic islands can usually be transferred thanks to the action of transposases and integrases (Dobrindt et al., 2004). Plasticity of bacterial genomes sometimes allows symbiotic islands to be recast into symbiotic plasmids, also called ICE (Integrative and Conjugative Elements; **Fig. 7c**) (Galibert et al., 2001; Haskett et al., 2016). Because of this organization of the symbiotic genes, it was proposed that the main driver of emergence and evolution of symbiosis is HGT (Poole et al., 2018). It seems that it is indeed the case as shown by the phylogeny of the symbiotic genes. Rather than clustering as their host genome (based on housekeeping genes), they group by host *biovar* (bv.) (Remigi et al., 2016; Rogel et al., 2011). This observation reflects the diversity of symbiotic couples and the host specificity (discussed later in this manuscript).

Usually, the symbiotic plasmids are dispensable in the free-living lifestyle of the rhizobia. In some cases, though, symbiotic plasmids are essential thanks to the transfer of essential functions from the chromosome into symbiotic plasmids. A good example of this process is the pSymB plasmid of *S. meliloti* on which the tRNA-Arg and *engA* (encoding a GTPase involved in ribosome biogenesis) genes were translocated. Due to this translocation, pSymB is essential both to symbiotic and free-living lifestyles, whereas pSymA is dispensable in the latter lifestyle (diCenzo et al., 2013; Poole et al., 2018).

... and what are legumes?

Legumes are plants from the *Fabaceae* family, which groups roughly 19500 species split in 765 different genera (Azani et al., 2017). The *Fabaceae* family, for which a modified classification was recently proposed, was initially split based on morphological features into three sub-families: *Caesalpinioideae*, *Mimosoideae* and *Papilionoideae* (Lewis et al., 2005). However, a study of the *matK* sequence phylogeny led to the dispersion of *Caesalpinioideae* legumes into different sub-families (**Fig. 8**). Four new sub-families were created: *Cercidoideae* (trees, shrubs & lianas), *Detarioideae* (trees & shrubs), *Duparquetioideae* (vines) and *Dialioideae* (trees & shrubs). The sub-family called *Caesalpinioideae* groups the remaining genera of the old

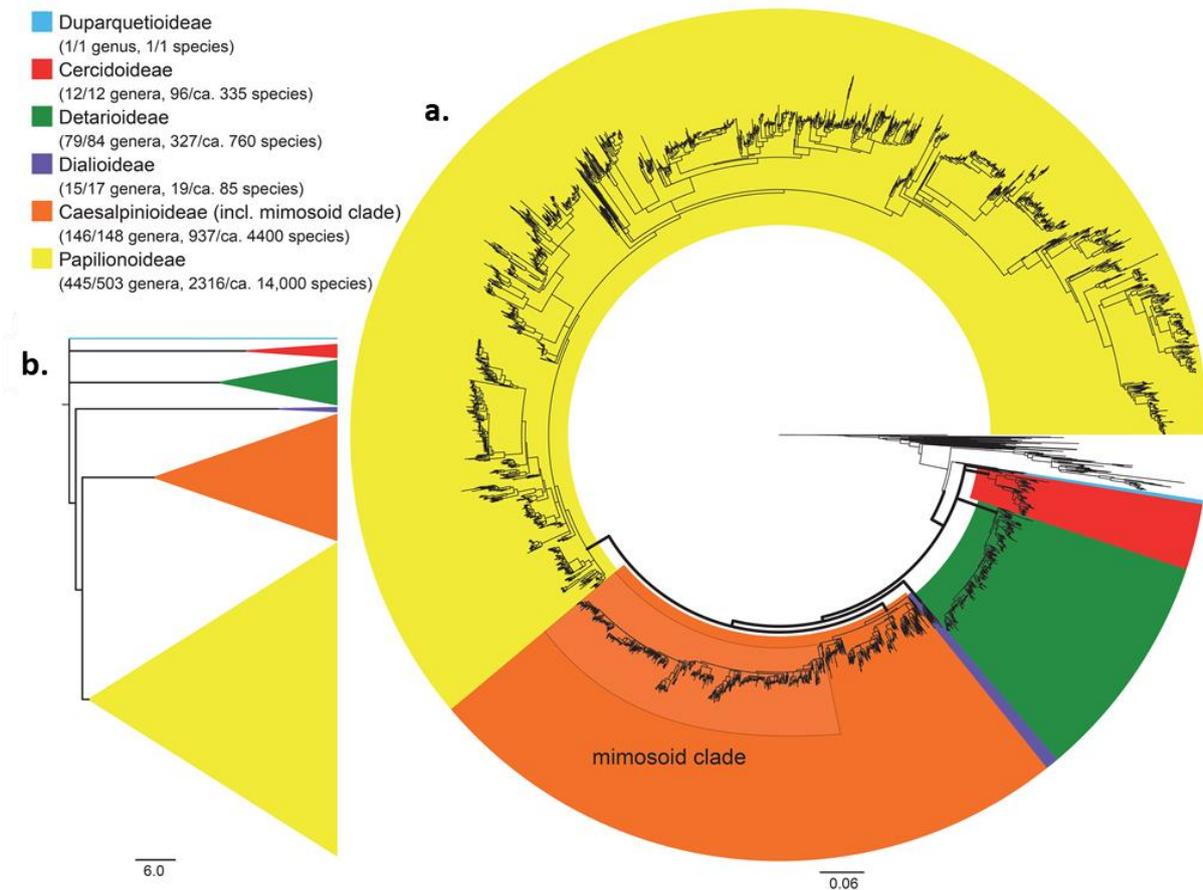


Figure 8 | Legume phylogeny. Phylogenetic trees from Azani & al. 2017 of the legume family built using *matK* sequences. These sequences represent 3696 of the ca. 19,500 species and 698 of the 765 genera. **a** | Phylogeny built with a Bayesian consensus method. **b** | Schematic phylogeny in which clade size are proportional to the number of species.

classification *Caesalpinioideae*, together with the *Mimosoideae* (Azani et al., 2017). The *Papilionoideae* is the most diverse sub-family of *Fabaceae* with up to 14 000 species identified (**Fig. 8b**) (Lewis et al., 2005). They can be found in a large range of environments (Azani et al., 2017) and are of great agronomical importance as many cultivated legumes are found in this sub-family. For example, legumes used either for human food production, such as bean (*Phaseolus vulgaris*), lentil (*Lens culinaris*), pea (*Pisum sativum*), and soybean (*Glycine max*), or as forage crops like alfalfa (*Medicago sativa*) and clover (*Trifolium repens*) are part of the *Papilionoideae* (Lewis et al., 2005). A major part of the *Papilionoideae* sub-family (97%) can interact with rhizobia to establish a nitrogen-fixing endosymbiosis (Bruneau et al., 2013). *Caesalpinioideae* also includes symbiotic host plants such as some acacias.

Within the *Papilionoideae* group, the Inverted Repeat Lacking Clade (IRLC) is the largest monophyletic clade. It is defined by a specific mutation: the loss of a 25 kb long region of the

chloroplast genome (Wojciechowski et al., 2000). Another group of interest is the *Dalbergioid* clade, which contains 180 different species originating either from America or sub-tropical regions of Asia or Africa. These plants can grow either on arid or semi-aquatic environments. These two clades display interesting symbiotic features that will be further detailed later in this manuscript.

Plants from the *Fabaceae* family were among the first cultivated species in the Fertile Crescent. Indeed, four out of the eight species that are considered as the founders of agriculture are legumes (Brown et al., 2009). Nowadays, soybean is the most cultivated legume in the world as it represents 65% (348 megatons) of world legume production in 2018 (FAO – <https://www.fao.org/faostat/>). This plant is ubiquitous in human and animal food. The ability of these legumes to perform symbiosis is of major interest in the ecology and agronomy fields. Indeed, thanks to this symbiosis, it is possible to grow these plants on marginal soils that are poor in nitrogen, without adding fertilizers. However, and somehow strikingly, the nodulating legumes are not the most efficient pioneer plants. Indeed, it seems that the ability to nodulate made them highly dependent on the soil microbiota, the presence of a compatible symbiont thus being necessary (Simonsen et al., 2017). Conversely, legumes which do not perform symbiosis are more efficient in pioneer colonization of environments. They are also more prone to become invasive species than the symbiotic ones (Simonsen et al., 2017). It may be possible that symbiotic legumes lost their efficiency to collect nitrogen in the soil, because the symbiosis relaxed this selection pressure. Therefore, the colonization of a new environment where compatible symbionts are not present yet is much more complicated.

General functioning and stakes around this biological process

The legume-rhizobium symbiosis is of the nutritional type, as opposed to protective or reproductive. Indeed, this symbiosis is initiated when plants lack nitrogen and leads to the formation of a symbiotic organ called nodule. This organ will host bacteria intracellularly, providing a controlled environment with nutrient supplies for the symbionts in exchange of fixed nitrogen. Thus, the plant can overcome the lack of nitrogen in the soil and thrive at the cost of energy and nutrients given to the bacteria.

Nitrogen is one of the four major components found in living organisms with oxygen, carbon, and hydrogen. It is for example found as a major constituent in macro-molecules such

as DNA or in proteins. This element is the most abundant component of our atmosphere, present at 78% under the form of dinitrogen (N_2). Paradoxically this nitrogen source is not usable by most organisms. Indeed, only some prokaryotes can reduce it to ammonium (NH_4^+) which can be used by most organisms. Only 0.001% of biosphere nitrogen can be assimilated by the majority of organisms, which makes it one of the most limiting elements in ecosystems (**Fig. 9**) (Newton, 1998). Most of the soil nitrogen is organic, coming from dead plants or animal corpses and feces. It is extracted from this matter and transformed into ammonia or nitrates by microorganisms (Miller & Cramer, 2004). However, in a crop this source of nitrogen is quickly depleted, and the addition of fertilizers is necessary to maintain a sufficient yield.

Although organic fertilizers were largely used since the advent of agriculture in the early Neolithic, an agricultural revolution took place in the beginning of the twentieth century. A chemical synthesis process was developed by Carl Bosch and Fritz Haber to produce ammonia from atmospheric nitrogen (Smil, 2011). With this process, named the Haber-Bosch process after its creators, 130 megatons of fertilizers are produced every year. This major advance propelled humanity in the green revolution and led to the drastic increase in human population, which became 4-times higher in the time of one century (Smil, 2011).

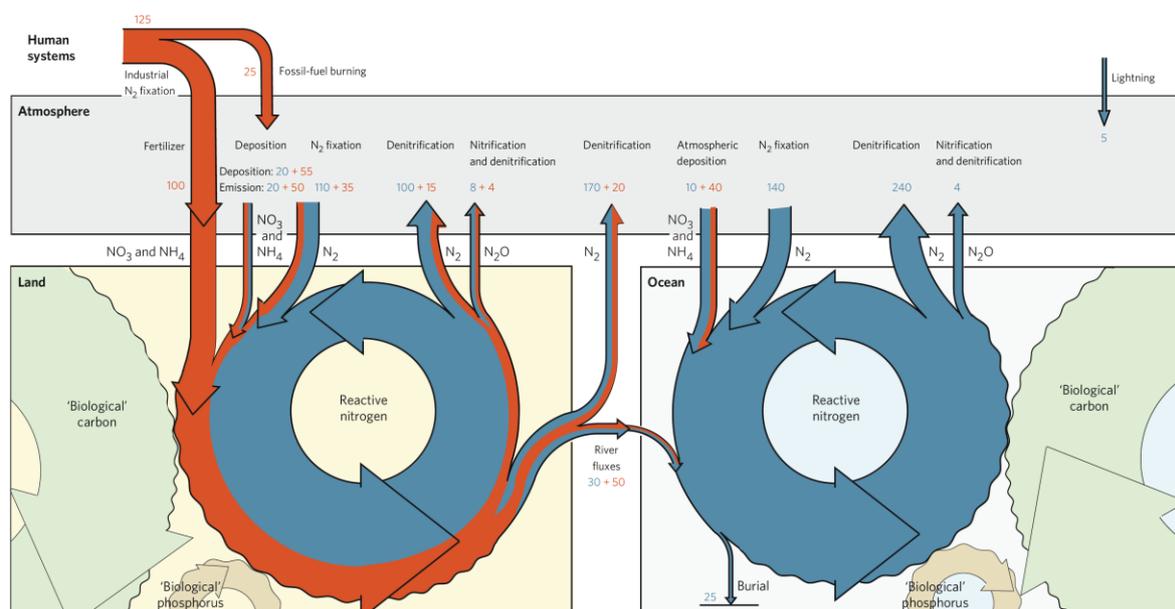


Figure 9 | Fluxes in the global nitrogen cycle on land and in the ocean. This graphic represents nitrogen fluxes in land (left panel) and ocean (right panel). The interaction with carbon (green) and phosphorus (orange) cycles is schematized. In the nitrogen cycle, blue fluxes represent “normal” fluxes whereas orange fluxes highlight disturbance by human activity. The numbers are expressed in Tg N per year. Note that N_2 fixation by diazotrophic bacteria is the major input of naturally fixed nitrogen in both land and ocean. Extracted from Gruber & Galloway 2008.

Nevertheless, this process has major drawbacks as well. It is highly energy consuming as it requires high temperatures (around 450°C) and high pressures (20-30 MPa). Moreover, application of nitrogen fertilizers is very harmful for the environment (Horrigan et al., 2002). Only 2-10% of nitrogen in chemical fertilizers would be effectively assimilated by plants (**Fig. 9**). The other fraction is generally either converted into gaseous forms or washed away by rains (Cao et al., 2018). This phenomenon increases the levels of nitrates in water, polluting groundwater and leading to the eutrophication of estuaries or ponds (Rastogi et al., 2015). Eutrophication designs the abnormal accumulation of nutrients and minerals in a body of water, which usually leads to the excessive development of algae. This last issue is a problem in aquatic ecosystem as it can deprive deep waters from oxygen such as in the Mississippi gulf (Fields, 2004). These are not the only drawbacks of the excessive utilization of synthetic fertilizers in crops. However, no serious alternative is available today to support the high food production required for sustaining the current world population.

Legume crops constitute only a small fraction of all cultivated plants (2-3 % - <https://www.fao.org/faostat/>), and it is therefore impossible to rely only on those plants and their biologically fixed nitrogen. Alternative cropping systems exist, such as bi-cultures, permacultures, fallow, or crop rotations, but these techniques are often efficient locally but not at a worldwide scale. Some of these techniques require either highly specialized equipment or a large workforce. Hence, there is a need to engineer new cultural techniques to overcome the negative issues arising from synthetic fertilizers. One approach is the use of nitrogen-fixing symbiosis as suggested already in 1917 by Burill & Hansen. Three different strategies can be established to reach this goal. The ability of fixing nitrogen could be transferred to plants themselves, expressing nitrogenase, the enzyme responsible for the nitrogen fixation, in specialized engineered organelles. Another option would be to add plant endophytic bacteria that fix nitrogen in crop cultures as a bio-inoculant. And the last strategy would be to transfer the ability to form symbiotic root nodules to cultivated plants such as cereals (Alunni & Mergaert, 2017). Each of these strategies has more or less critical milestones that need to be solved in order to succeed. However, these strategies place in the legume-rhizobium symbiosis the hope of a new sustainable and environment-friendly agriculture.

ii. The initiation of the symbiosis – A molecular dialog between plant and bacteria

To set up this interaction, both partners exchange information by the mean of signal molecules. These molecules initiate changes in the physiology of the receiver. This exchange happens as a two-step dialog first involving flavonoids secreted by the plant and then Nod-factors secreted by the bacteria. Initiation of this symbiosis shares features with mycorrhization as plants also secrete flavonoids and strigolactones to attract fungi. The potential symbionts respond with Myc-factors, similar in structure to Nod-Factors (Maillet et al., 2011).

The plant request – Flavonoids as a “message in a bottle”

When in condition of nitrogen starvation, legume plants synthesize and secrete a mixture of phenolic secondary metabolites known as flavonoids (**Fig. 10a**). This family of molecules have many roles in plants and can be produced in different organs for various purposes, ranging from flower pigmentation to protection against pathogens. However, the ones that are secreted through the roots are addressed to potential symbionts in the soil.

Some bacteria, and more precisely rhizobia can be able to recognize these molecules, provided they have a compatible receptor. Flavonoids, after being internalized in the bacterium, bind to a protein, usually NodD1 or NodVW. In the case of NodD1, it is a constitutively expressed transcription factor, which becomes activated probably by conformational change after binding a compatible flavonoid. NodVW is a two-component system, involving a bacterial membrane receptor (histidine kinase) for signal transduction and a response regulator that activates transcription of target genes when it is phosphorylated by the flavonoid-bound receptor. Both recognition systems lead to the expression of the *nod* genes (**Fig. 10b**) (N. K. Peters et al., 1986). Simultaneously, flavonoids will act as chemo-attractants, inducing the responsive bacteria to migrate along the flavonoid gradient and towards the root apparatus of the legume (Dharmatilake & Bauer, 1992; Loh & Stacey, 2003; Phillips & Tsai, 1992).

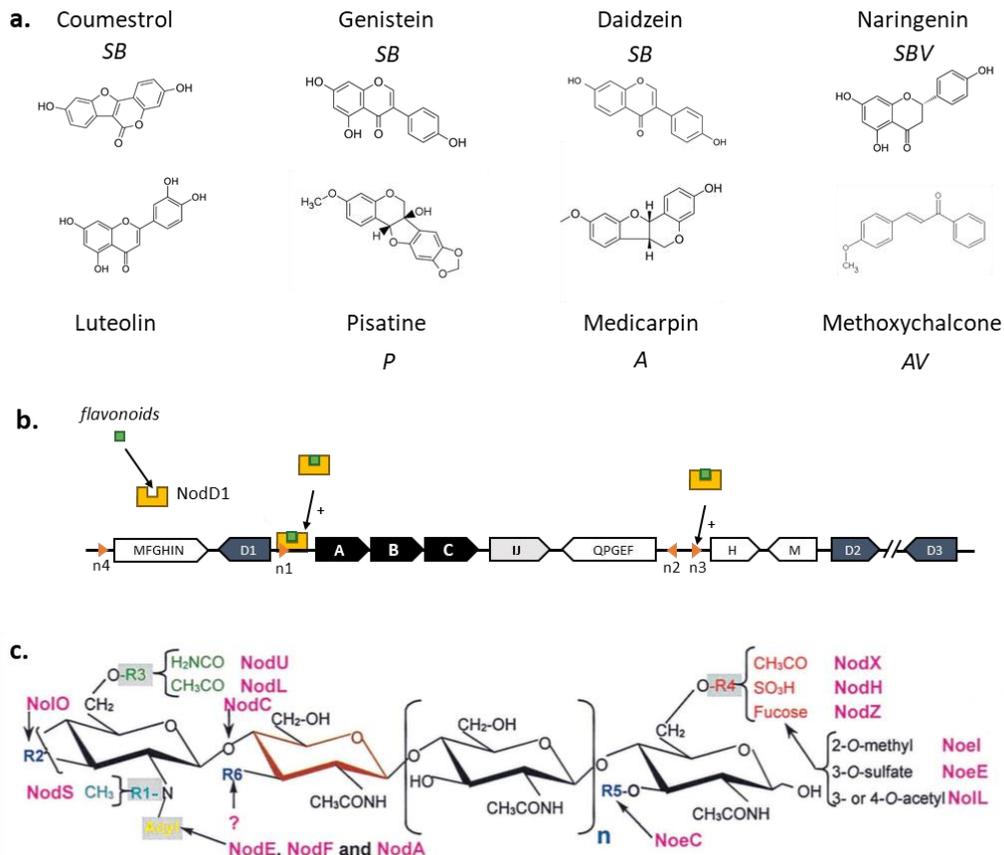


Figure 10 | Molecules involved in the molecular dialog. Examples of flavonoids (**a**) and Nod-factors (**c**) structures along with the activation of the *nod* genes by NodD1 and flavonoids (**b**). **a** | Flavonoids shown to be involved in the induction of *nod* genes in rhizobia. Capital letters represent for each flavonoids the plant genera in which it can be found. A: alfalfa; B: bean; P: pea; S: soybean; V: vetch. Luteolin is not specific of legume plants but has been shown to induce *nod* genes as well. **b** | In *Sinorhizobium meliloti*, when flavonoids interact with NodD1, the complex will bind to *nod* cassettes (orange arrows annotated n1 to n4). This binding activates the expression of the *nod* genes whether they are canonical (*nodABC* in black) or strain-specific (other colors). Adapted from Cren et al., 1995. **c** | Nod-factor structure with possible substitutions identified in different rhizobia. Genes responsible for these modifications are in pink font. R2: OH, carbonyl; R5: arabiosyl; R6: OH, acetyl and fructosyl.

The bacteria answer – Nod Factors as the trigger

If the bacteria can recognize the flavonoids, NodD will be activated. This factor can bind a consensus sequence called *nod* box located in the promoter sequence of the *nod*, *noI* and *noE* target genes. These genes are involved in the synthesis of a molecule that will activate plant responses allowing bacterial infection and nodule organogenesis to occur. This signal molecule, called the Nod factor, is a lipo-chito-oligosaccharide (LCO; **Fig. 10c**). Nod factors are first synthesized by the enzymes encoded by the *nodABC* genes (Mergaert et al., 1997). These genes are present in every nodulating bacteria except in some photosynthetic *Bradyrhizobium* strains such as ORS278 and BTAi1 (Giraud et al., 2007). They produce the core LCO, which will

then be decorated by several modifications to produce the final Nod factor molecules. It is not one type of LCO that is secreted by a bacterium but a cocktail of LCOs with varying structures. A large set of genes can be involved in this process as it has been demonstrated in the strain *Sinorhizobium fredii* NGR234 where thirteen genes encode enzymes that modify the core LCO (Broughton et al., 2000). These modifications are diverse ranging from methylation, acetylation, carbamylation, sulfation to glycosylation (**Fig. 10c**). Once Nod factors are assembled and decorated, a cocktail of these molecules will be secreted by the NodIJ ABC (ATP-Binding Cassette) transporter in the rhizosphere (D'Haese & Holsters, 2002; Lerouge et al., 1990; Mergaert et al., 1997).

If the plant can detect the Nod factor, *id est* if it has the compatible receptors, the initiation of the symbiosis will pursue (**Fig. 11a**). The plant recognizes Nod factors using LysM-RLK (Receptor-Like Kinase) type receptors. These trans-membrane proteins are called NFP/LYK3 in *Medicago truncatula* or NFR1/NFR5 in *Lotus japonicus* and form heterodimers. These receptors function together with DMI2/SymRK, an LRR-RLK family receptor (Leucine Rich Repeat) (Limpens et al., 2005; Oldroyd, 2013). Similarly, initiation of AM symbiosis involves LysM-RLK and SymRK to perceive the Myc-Factors (Chiu & Paszkowski, 2020).

This recognition will activate a complex signaling cascade leading to the induction of symbiotic genes (**Fig. 11b**) (Zipfel & Oldroyd, 2017). This cascade was discovered through genetic screens and the identification of plant mutants unable to initiate a functional symbiosis. Two main species were involved in these studies: *M. truncatula* and *L. japonicus*. Recognition of Nod-Factors by plant receptors leads to the initiation of calcium spiking at the nuclear envelope. However, how the calcium spiking is induced is not yet fully understood. However, it has been suggested that mevalonate would be an intermediate between the signal perceived at the plasma membrane and the calcium spiking at the nucleus envelope (Venkateshwaran et al., 2015). This hypothesis arises from two observations. First, an enzyme from the mevalonate biosynthesis pathway, HMGR1 (3-Hydroxy-3-Methylglutaryl CoA Reductase 1), directly interacts with DMI2 that is part of the Nod factor receptor complex (Kevei et al., 2007). Second, exogenous application of mevalonate is enough to induce calcium spiking on nuclear envelopes through DMI1 (Venkateshwaran et al., 2015). The calcium spiking involves many proteins located on the nuclear envelope. The ion-channel DMI1 in *M. truncatula*, or its

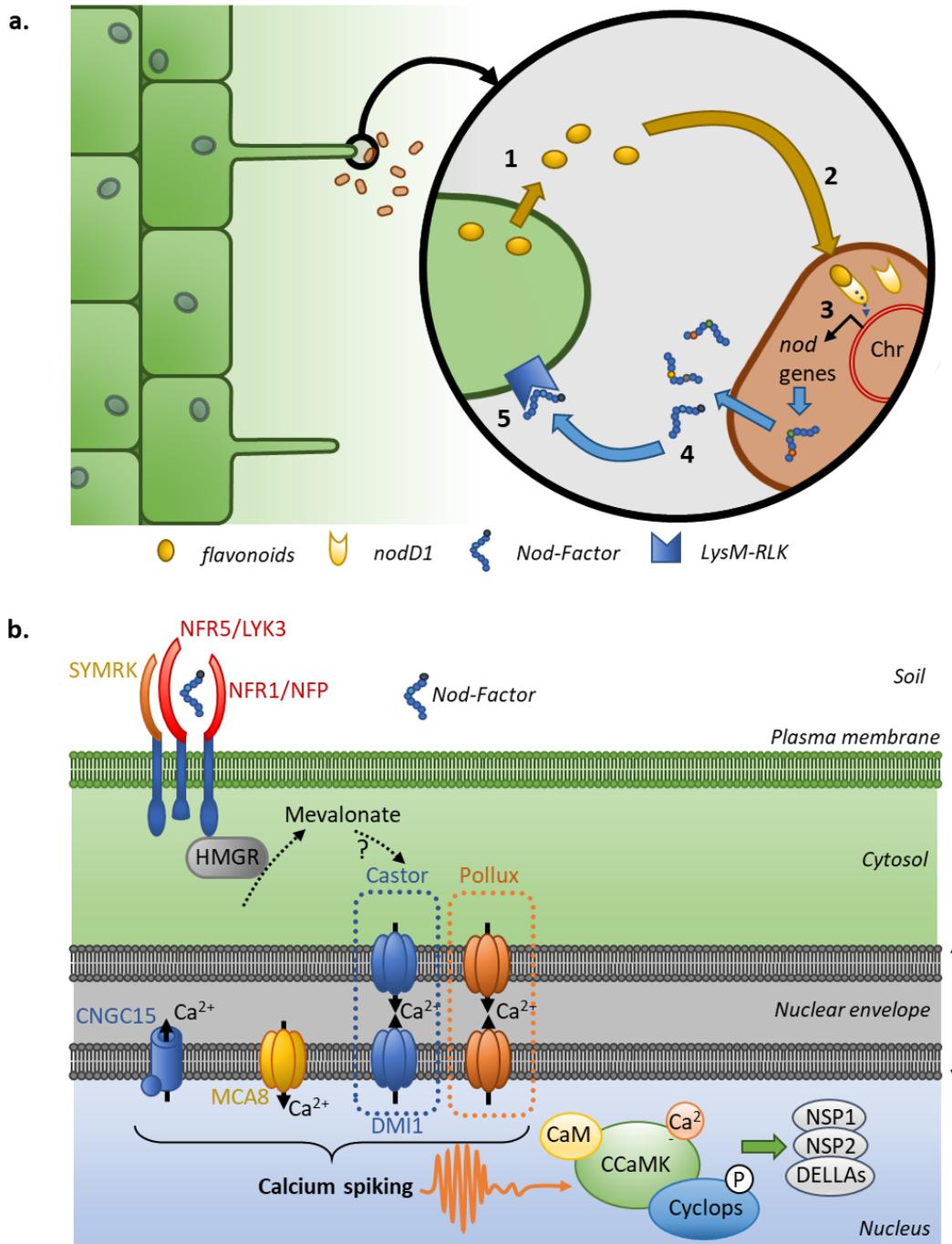


Figure 11 | Molecular dialog between legume and rhizobia. This scheme depicts the different steps of the molecular dialog as well as its consequences for plant cells. **a** | Flavonoids are secreted by roots cells in the soil to initiate a symbiotic interaction with rhizobia (1). A compatible symbiont can perceive them and synthesize Nod-Factors (3) that can be perceived by the legume plant receptors after secretion (4-5). In this scheme, the nod factor receptor presented is NodD1. NodVW is located on the bacterial envelope. **b** | Activation of symbiotic genes subsequent to Nod-Factor perception. A complex signaling cascade is triggered involving LysM-RLK and LRR-RLK membrane receptors that activate mevalonate signaling and a set of nuclear proteins triggering calcium spiking in the nucleus. Finally, the calcium signal is decoded by and translated into the activation of target gene expression. Symbiotic genes are responsible for two concomitant processes: organogenesis and infection.

homologs CASTOR & POLLUX in *L. japonicus* are involved in this calcium spiking (Ané et al., 2004; Charpentier et al., 2016). These channels are Ca^{2+} transporters that regulate the activity of a Ca^{2+} channel called CNGC15 (Kim et al., 2019). This channel imports Ca^{2+} into the nucleus, whereas another transporter called MCA8 exports it (Capoen et al., 2011; Charpentier et al., 2016; Zipfel & Oldroyd, 2017). It is believed these calcium transporters, by inducing successive membrane polarization and depolarization, are responsible for the calcium spiking (**Fig. 11b**).

Once the calcium spiking is ongoing, new actors get involved, this time to decode this signal. The first factor is called CCaMK (for Calcium/Calmodulin dependent protein Kinase) or DMI3 (Lévy et al., 2004). When this protein detects oscillations in the levels of calcium in the nucleus, CCaMK is able to phosphorylate a factor called CYCLOPS (Yano et al., 2008). Once activated, CYCLOPS becomes able to recruit transcription factors such as DELLA, NSP1 and NSP2. These transcription activators induce the expression of downstream symbiotic genes (**Fig. 11b**) (Zipfel & Oldroyd, 2017).

The target genes of DELLA, NSP1 and NSP2 coordinate the simultaneous activation of two processes. On one hand, these genes control the infection process, during which bacteria colonize the plant tissues until the intracellular infection of plant cells. On the other hand, an organogenesis program is initiated, leading to the formation of a nodule primordium that will grow out to mature, infected nodules.

Remarkably, a portion of this signaling pathway is shared between the mycorrhizal and legume-rhizobium symbioses and is therefore known as the Common Symbiosis Signaling Pathway (CSSP) (Oldroyd, 2013). Although being triggered by different types of LCOs and LysM receptors, the CSSP is triggered in both cases and the signaling cascade leads to calcium spiking followed by the activation of the CCaMK/CYCLOPS complex and some common transcription factors (Chiu & Paszkowski, 2020). Nevertheless, in the end, different sets of transcription factors are recruited, and different transcriptional responses are activated during the two symbioses. It is unclear how the plant can select different transcription factors from a common signaling pathway, but some unknown signals might be responsible for inducing different responses.

The notion of host spectrum

Even though the flavonoid and Nod factor signals, as well as the CSSP, are conserved in most legume-rhizobium interactions, there are nevertheless high chances that the molecular dialog leads nowhere. Indeed, the recognition of either flavonoids by the bacteria or Nod factors by the plant is critical. Depending both on the molecular nature of the “messengers” and the compatibility of the receptor, the interaction can be initiated or not. This feature explains why some plants or bacteria have a limited host range. One good example is *Azorhizobium caulinodans*, which only associates with few species from the *Sesbania* genus (Sprent et al., 2017). In some cases, specialization can be observed within a single species and different biovars (bv.) emerged which each associate with specific host plants. *Rhizobium leguminosarum* behaved that way, and today different biovars exist (such as *phaseoli*, *trifolii* and *viciae*). Each of them interacts with the plant genera they are named after (Rogel et al., 2011). On the opposite, the strain *S. fredii*NGR234 can associate with 112 legume genera based on our current knowledge. On the plant side, *Sophora flavescens* can associate with no less than 35 different bacterial species (Jiao et al., 2015) while other plants, like *Medicago* spp. accept a much more restricted range of rhizobial species. Several other examples have similar characteristics (S. Goormachtig et al., 1998; Krysciak et al., 2014; Poole et al., 2018; Pueppke & Broughton, 1999; Zhao et al., 2010).

Eliciting nodule formation – Nod-dependent and Nod-independent symbioses

In the last decade, the study of tropical symbioses revealed atypical processes to initiate the interaction. It was discovered that two symbiotic photosynthetic *Bradyrhizobium* strains, ORS278 and BTAi1, lack the *nodABC* genes implying that no Nod factor can be synthesized by these strains. Strikingly, they are still able to perform symbiosis with some *Aeschynomene* species, belonging to the so-called group III (Chaintreuil et al., 2013; Giraud et al., 2007). This group is part of a classification of *Aeschynomene* species based on cross-inoculations of 26 different *Aeschynomene* species and 15 *Bradyrhizobium* strains (Alazard, 1985). In the group III, plants are the only ones that can perform symbiosis with bacteria that lack Nod factors. Tellingly, *Bradyrhizobium* sp. ORS285 loses its ability to nodulate *Aeschynomene* species from

group I when *nodB* is deleted but can still induce functional symbiosis with *Aeschynomene* from group III (Giraud et al., 2007).

More recently, in 2013, another type of Nod factor-independent symbiosis has been discovered. Studies of the *G. max* var. "Enrei" – *Bradyrhizobium elkanii* USDA61 interaction showed that a symbiont lacking Nod factor production through the deletion of *nodC*, could still nodulate this host plant, even if this one lacked the Nod factor receptor NFR1 (Okazaki et al., 2013b). In this specific condition, nodulation initiation relies on the presence of a functional Type 3 Secretion System (T3SS) in the bacterium. The T3SS is a quite common feature in bacteria interacting with plant and animal hosts. It allows the secretion of bacterial proteins called type three effectors into the host cytoplasm. Usually, the bacteria uses it to shut down host defense reactions and promote host colonization, both in pathogenic and mutualistic interactions. When the T3SS of *B. elkanii* USDA61 was mutated in addition to *nodC*, the bacteria would lose their ability to perform symbiosis with soybean (Okazaki et al., 2013b). It was concluded that one or several type III effectors somehow re-activate the symbiotic signaling pathway downstream of the Nod factor receptor. Other non-photosynthetic bradyrhizobia were subsequently found to use their T3SS instead of Nod factors to nodulate group III *Aeschynomene* (Okazaki et al., 2016). Conversely, the majority of photosynthetic *Bradyrhizobium* species do not have a T3SS. Three types of symbiotic signaling pathways can thus be distinguished: Nod factor-dependent, Nod factor-independent and T3SS-dependent, or Nod factor-independent and T3SS-independent signaling (Okazaki et al., 2013b). Although using different mechanisms, these three types of symbiosis initiation seem to converge towards the activation of similar plant responses.

The molecular mechanisms that underlie these Nod factor-independent symbioses are still poorly known today (Fabre et al., 2015), but this feature raises interesting questions on the evolution of symbiosis. Further investigation might unveil whether Nod factor-independent symbioses were anterior to Nod factor-dependent symbioses or *vice versa*. As only the *Aeschynomene* species from group III are naturally associated with Nod factor-independent symbioses and as this group is monophyletic, this feature could have emerged recently and specifically for these species. On the other hand, Nod factor independent symbiosis is possible in particular conditions as in the above-described soybean – *Bradyrhizobium elkanii* interaction

or in the *L. japonicus* – *Mesorhizobium loti* interaction. In the latter interaction, it was shown that even if plant Nod factor receptors were mutated as well as bacterial *nod* genes, the symbiosis could still happen with very weak efficiency (Madsen et al., 2010). This observation was taken as an argument for the ancestral status of Nod factor-independent symbiosis. Thus, two opposite conclusions have been reached concerning the evolutionary status of Nod factor-independent symbiosis. Clearly, additional work is required to settle this question.

iii. The initiation of the symbiosis – An infection process leading to endosymbiosis

As already mentioned, once the molecular dialog between legume and rhizobia is set up, two processes are initiated concomitantly: the infection of the plant cells by the symbionts, and the organogenesis program that will lead to the formation of the symbiotic nodule.

The different stages of the infection – Shepherd's crook model

Beginning with the infection process, bacteria are attracted by chemotaxis using a flavonoid gradient to migrate towards the roots. When they reach plant root hairs, they will bind to their apex. Attachment of the bacteria to the cell wall of the root epidermis occurs in two steps with plant lectins that bind to bacterial exopolysaccharides, followed by the secretion of rhizobial proteins known as rhizadhesins that reinforce the binding (Matthysse & McMahan, 1998; Smit et al., 1991). Eventually, cellulose fibrils are synthesized by the bacteria, helping to consolidate the structure (Ausmees et al., 1999). Although these adhesion steps have been detailed in some interactions where they seem to be essential, it is still unclear to what extent this process is conserved in other legume-rhizobium interactions. While this adhesion is carried on, the recognition of Nod factors by the root hair induces its curling, resulting in a root hair cell with a characteristic shape that got the name of "the Shepherd's crook" (**Fig. 12a**). The root hair curling takes about two to four hours. The small space that is formed in the inner region of the hook, the infection pocket, will enclose a few bacteria that will enter the plant tissues and be the founder cells of the nodule bacterial population.

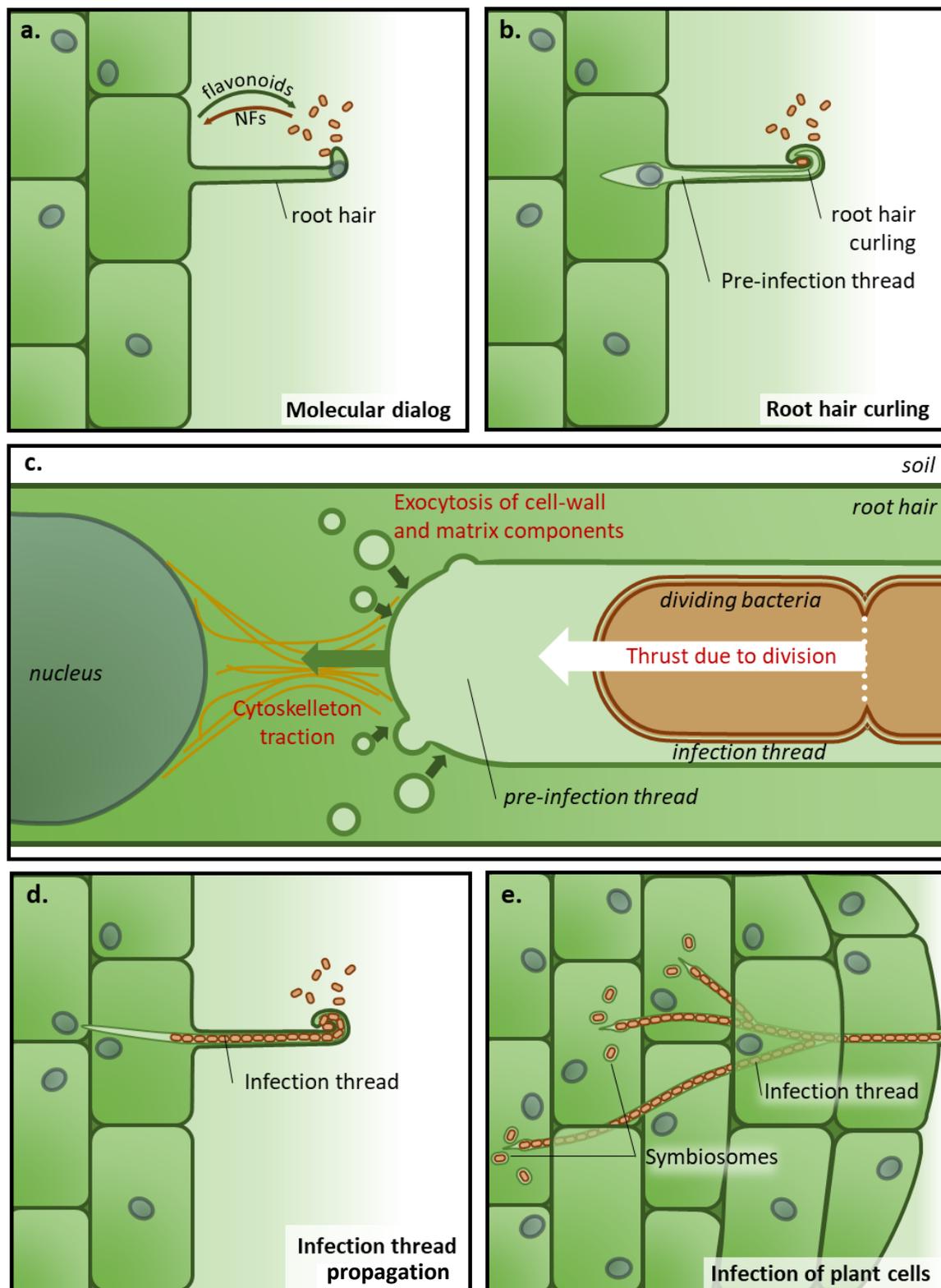


Figure 12 | Infection of legume roots by rhizobia. Once the molecular dialog is successful (a), the tip of the root hair begins to curl. Bacteria will develop in the hook and eventually break through plant cell wall. Thereafter, a pre-infection thread will begin to form by remodeling the cytoskeleton and the endomembrane system (b), helping the progression of bacteria through the root hair. Several processes add up to build the infection thread and help the migration of bacteria (c, d). In the end, bacteria will be released into host symbiotic cells of the nodule (e). They form symbiosomes and are the first symbionts of the developing nodule. Adapted from Oldroyd 2013 and Oldroyd & Dixon, 2014.

In the infection pocket, bacterial and plant cell wall digesting enzymes (cellulases, polygalacturonases, pectate-lyases) are secreted (Gage, 2004; Ramos et al., 2008; Xie et al., 2012). The action of these enzymes will initiate the penetration of bacteria in the root hair cell, ten to twenty hours after contact. In this small niche created by the plant, bacteria will begin to multiply, while the cytoskeleton of the plant cell is rearranged (**Fig. 12b**). This change in the cytoskeleton is coupled with exocytosis of matrix and cell-wall compounds that helps the formation of the infection thread (IT) (Oldroyd et al., 2011). Moreover, a pre-IT is formed by the root hair nucleus that is dragged down to the cell base by the cytoskeleton. These two forces combined help the formation of the IT (**Fig. 12b&c**) (Fournier et al., 2008). This IT is the channel by which bacteria enter plant tissues. The IT progresses through the root hair until it reaches its basis while bacteria migrate through this structure using the mechanical force generated by their multiplication (Gage, 2002). To reach the cortical cells the pre-IT propagates in the underlying cell layers. To do so, the nuclei of the cells that will be crossed by the IT align and the cells form cytoplasmic bridges. When an IT reaches a target cell, the cell wall is locally degraded, the membrane invaginates and the tube structure of the IT propagates in the new cell (**Fig. 12d&e**) (Oldroyd et al., 2011). When IT and their bacteria reach the cortical cells of the root, they will infect the first symbiotic cells intracellularly (**Fig. 12e**). Bacteria become bacteroids and they penetrate the inside of plant cells thanks to an invagination of the membrane, similarly to endocytosis, thus resulting in a particular structure called symbiosomes. This structure is composed of the bacteroid wrapped in the plasma membrane of the plant cell. The space between bacteroid and the symbiosome or peri-bacteroid membrane is called peri-bacteroid space (Jones et al., 2007).

Infection processes in Aeschynomeneid species.

The root hair infection mechanism that is described above is found in the majority of contemporary legumes, but it is not the only possible way for infecting plants during symbiosis. In some tropical legumes, as for example in the Dalbergioids, bacteria can enter plant tissues via lateral root bases.

Two scenarios are known for bacteria to infect *Aeschynomene* plants (**Fig. 13**) (Bonaldi et al., 2011). In *Aeschynomene indica*, the base of axillary root hairs located at the basis of lateral roots is usually colonized by bacteria (**Fig. 13a**). These bacteria constitute the first

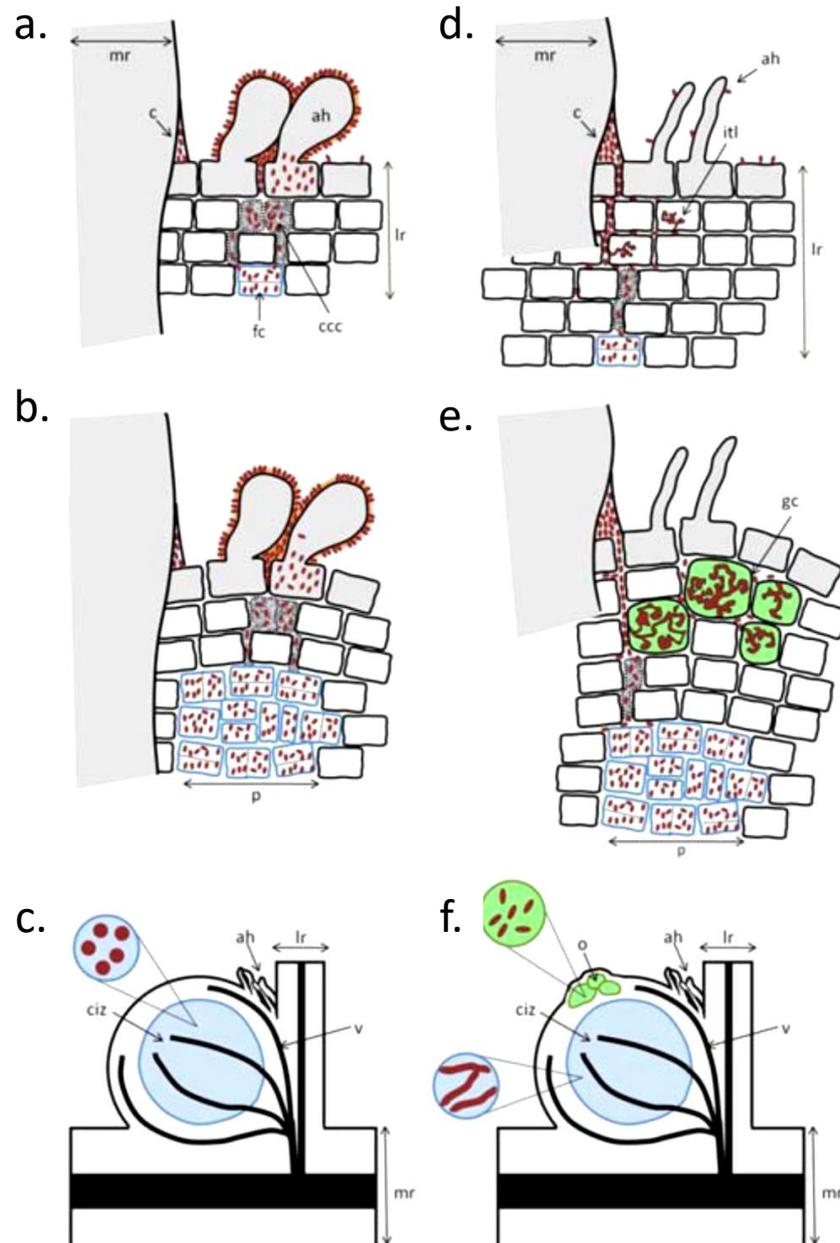


Figure 13 | Infection of *Aeschynomene* species roots by rhizobia. These schemes compare the infection of *Aeschynomene indica* (a through c) and *Aeschynomene afraspera* (d through f). a & d represent plants at 1 to 2 days post inoculation (dpi) whereas b & e represent plants at 3 & 4 dpi. c & f depicts 14 dpi mature nodules. *Aeschynomene indica* symbiosis begins with the infection of cortical cells by the bacteria present on lateral root hairs (a). The cortical cells are lysed and collapse, leading to the infection of a deeper cortical cell. This cell, called the founder cell will be at the origin of the nodule (b). Mature nodules host spherical bacteroids (c). In the plant *Aeschynomene afraspera*, bacteria enter root tissues through the cracks due to the emergence of lateral roots (d). After migrating to the lower layers of cortical cells, they infect the founder cell (e). On their way, they infect cortical cells of higher tissue level, which will form the outgrowth (f). Mature nodules formed from the founder cells host elongated bacteroids. On their tip is placed the outgrowth, which host undifferentiated bacteria. Abbreviations: mr, main root; lr, lateral root; c, crack; ah, axillary root hairs; ccc, collapsed cortical cells; p, primordium; fc, founder cell of the nodular primordium; gc, giant cell; o, outgrowth; itl, infection-thread like structures, v, vascularization; ciz, central infected zone. Epidermal cells (gray) and cortical cells (white). Extracted from Bonaldi et al. 2011.

inoculum for the symbiotic interaction. Bacteria will penetrate the apoplast of the root through the interstices near root hairs, reaching the first layer of cortical cells. There, they will multiply in the space generated by the death of the nearby cortical cells (**Fig. 13a**). Afterwards, this space collapses under the turgor pressure from adjacent cells, thus pushing bacteria deeper into the root cortex. Bacteria then infect the nearby cortical cells giving birth to the founder cells which will be at the origin of the future nodule (**Fig. 13b&c**) (Bonaldi, et al., 2011). On the other hand, in *Aeschynomene afraspera*, axillary root hairs, which are longer and sparsely distributed on the root compared to *A. indica*, are not as extensively colonized by bacteria as in *A. indica* (**Fig. 13d**). Thus, rather than an infection via these root hairs, bacteria penetrate root tissues via wounds called cracks that form during lateral root formation (**Fig. 13d**). This process is called infection by “crack-entry” or lateral root base nodulation. It can also be found in plants like *Sesbania rostrata* or *Arachis hypogaea*. Bacteria will migrate through the cortex to infect the first founder cells of the nodule, three to four cell layers deep in the root (**Fig. 13d**). These infected founder cells, as in *A. indica*, will divide leading to the formation of the organ composed of infected cells. On their way to the inner cortical cells, bacteria will also infect some outer cortical cells thanks to infection thread-like structures (**Fig. 13e**). These infected cells will become giant and will form a structure called outgrowth at the tip of the nodule, which host a pool of non-fixing bacteria (**Fig. 13f**).

It is worth noting that in those two cases, no infection threads nor induction of a distant primordium is observed (Bonaldi et al 2011). In the case of *A. afraspera*, this is striking, as Nod factor dependent infection is usually associated with these features. In *A. indica* and *A. afraspera*, nodules can also form on the shoots. Although described in less detail, the infection proceeds with the crack-entry process (Chaintreuil et al., 2013).

How is the infection regulated?

Besides Nod factor signaling, both partners use several additional functions to ensure a smooth progress of the infection (Jones et al., 2007). One of them is the presence of coating polysaccharides (LPS, EPS, KPS and GC) on the surface of the bacteria, which can be recognized by plant receptors and then tune negatively or positively plant defense responses. For example, KPS seems to have various roles, ranging from the stimulation of flavonoid production in the *S. meliloti* – alfalfa interaction, to an involvement in the efficiency of symbiosis (Becquart-de

Kozak et al., 1997; Kereszt et al., 1998; Le Quéré et al., 2006). However, it is unsure whether these roles are really linked to KPS only as their biosynthesis pathways are partially common with those of the EPS.

EPS were shown to be essential actors at this step of the infection, either required to properly infect plant tissues as in pea or alfalfa (Breedveld et al., 1993; Cheng & Walker, 1998) or helping the plant to recognize the bacteria as a non-aggressive invader (Gourion et al., 2015; Gully et al., 2016; Kawaharada et al., 2015; Lehman & Long, 2013; Zipfel & Oldroyd, 2017). This last example, which was demonstrated in the *L. japonicus* – *M. loti* interaction (Kawaharada et al., 2015), involves a transmembrane LysM-RLK receptor called EPR3. This receptor, expressed after Nod factor perception, recognizes EPS, and acts differently depending on the EPS structure it detects. No EPS recognition slightly affects the efficiency of symbiosis compared to the recognition of a compatible EPS by EPR3. However, if EPR3 binds to an incompatible (truncated or modified) EPS, it leads to the cessation of the infection process. This feature is depicting very well the multiple checkpoints that are active all along the symbiotic process to maximize the chances of success, reducing the risks of colonization by pathogenic or selfish strains. Other surface polysaccharides play part in this process, such as CGs, which might act as signal molecules. CGs are important for the infection, though the underlying mechanisms are not unveiled yet (Ciocchini et al., 2007; Crespo-Rivas et al., 2009).

iv. The initiation of the symbiosis – Nodule formation and structure

Indeterminate nodules of the Medicago genus

While bacteria infect the first cortical cells, the nodule begins to develop. Different nodule structures exist, specific to certain legumes genera, but in this section the nodule structure of *Medicago* species will be described. The primordium of the nodule, born from cortical cells of the root, hosts the first bacteria and begins to divide, thus forming the nodule. At the tip of the developing nodule is the meristem positioned, responsible for the nodule growth. These nodules are of the indeterminate type because the meristem will be maintained all along the nodule lifespan (**Fig. 14b**).

Three layers of tissues form the peripheral zone of the nodule. First, the nodule cortex, which acts as a physical barrier and structural support for the symbiotic organ. This tissue

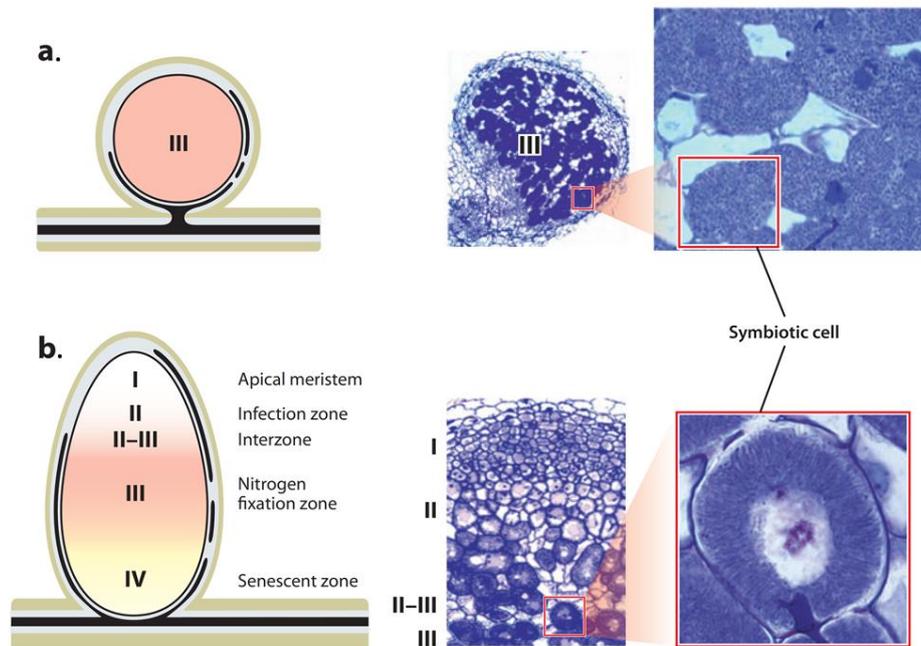


Figure 14 | Comparing determinate and indeterminate nodules. This figure presents both determinate (a) and indeterminate (b) nodule structures. Both panels display from left to right: a schematic of the nodule structure, a longitudinal nodule section after toluidine blue staining, and an enlargement of a symbiotic cell. Extracted from Kondorosi et al., 2013.

envelops the endodermis, which limits oxygen diffusion in the nodule tissues. Inside this envelope lies the vascular system, developed from the roots. Unlike root vasculature, which is central, nodule vascular bundles surround the central part of the nodule where symbiotic cells host the bacteroids, thus favoring the metabolic exchanges between bacteroids and the rest of the plant (Kondorosi et al., 2013).

In the central part, indeterminate nodules can be dissected into different functional parts, organized in layers from meristem to root (Vasse et al., 1990). We can count five distinct zones which are, from tip to root: the meristem (I), the infection zone (II), the intermediate zone (II-III), the fixing zone (III), and the senescence zone (IV) (**Fig. 14b**) (Kondorosi et al., 2013). The meristem is composed of actively dividing cells, driving the development of the nodule. The infection zone is the part of the nodule where cells derived from the meristem undergo a differentiation process and that contains dividing bacteria within infection threads before their release into plant cells. After their release in the plant cells, bacteroids enter the intermediate zone where they begin their physiological transition to become nitrogen-fixing. This state is reached in the fixing zone where bacteroids efficiently fix the atmospheric dinitrogen into

ammonia. The last layer, which appears in nodules as they grow older, is the senescence zone which will be further described later in this section.

Plant cells of the nodule are usually polyploid in the fixing zone. This feature begins to appear in the infection zone, initiated when bacteria infect plant cells. Indeed, meristematic cells are usually diploid, codified by 2C, 1C referring to the size of the haploid genome, or have two times their DNA content (4C) after genome replication at the onset of mitosis. In the infection zone cells can have up to 8-16C. These numbers increase as plant cells reach the fixing zone where they can have up to 64C (Cebolla et al., 1999; González-Sama et al., 2006; Vinardell et al., 2003). This endoreduplication of plant cells seems necessary for a functional symbiosis as shown by the study of a plant cell-cycle regulator, Ccs52A. This protein is an activator of the anaphase-promoting complex involved in the degradation of cyclins. Cyclins are central regulators of the cell-cycle together with CDK (Cyclin-Dependent Kinases), thus promoting cell cycle progression. Ccs52A is required for the endoreduplication and the presence of a fixing zone in the nodule. Indeed, when its levels are reduced by RNAi, the structure of the nodule is altered with a prominent senescence zone and an absence of fixation (Cebolla et al., 1999; González-Sama et al., 2006; Vinardell et al., 2003).

Comparing indeterminate and determinate symbiotic organs

As described before, in the *Medicago* - *Sinorhizobium* interaction, nodules are indeterminate. This means they will keep growing over time thanks to the persistence of the meristematic zone during all their lifespan, the nodule eventually reaching a senescence step (see part *viii* of this section). This kind of nodules can usually be found in the legume sub-family called *Faboideae*. For example, in this sub-family, indeterminate nodules are found in plants belonging to the genera *Medicago* (alfalfa), *Pisum* (pea), *Trifolium* (clover) or *Vicia* (vetch). Other legumes form indeterminate nodules such as acacias or *Sesbania* plants. Interestingly, *Sesbania* is also able to form determinate nodules, depending on culture conditions (Fernández-López et al., 1998).

The other type, the determinate type of nodules structure is quite different from the indeterminate nodules as the cell layers that can be found in indeterminate nodules are absent (**Fig. 14a**). Contrary to indeterminate nodules, their bacteria-hosting region is homogeneous.

Determinate nodule development is a sequential process as compared to the spatial development observed in indeterminate nodules. Thus, determinate nodules have no persistent meristem and as a consequence, their growth is determined in time. These nodules have a typical spherical shape as compared to the more elongated indetermined nodules. Symbiotic cell development and bacteroid differentiation are also synchronized over the whole central tissue (Kondorosi et al., 2013). Peripheral tissues of determinate nodule are composed of a parenchyma in which the vascular system is embedded, the endodermis and in most outer part the cortex (Hirsch, 1992; Kondorosi et al., 2013). Eventually senescence of the nodule will begin by the center of the fixing zone and will spread radially until the whole nodule is senescent. These determinate nodules can be found in many legumes including *Glycine* (soybean), *Lotus* (trefoil), *Phaseolus* (bean), *Vigna* (mung bean) among others.

v. The fixing stage of the process and beyond

Nitrogenase – The enzymatic complex responsible for nitrogen fixation

When nodules mature, they host bacteroids that actively fix nitrogen and transfer it to the plant. The enzymatic complex responsible for nitrogen fixation is called nitrogenase and, in the context of symbiosis, it is specifically expressed in the fixation zone. This complex is not specific to plant symbionts as other prokaryotes, collectively called diazotrophs, possess nitrogenase and are able to fix nitrogen in a free-living state, *id est* not associated with a host. Nitrogenase catalyzes the following reaction:



This formula points out the high levels of energy and reducing power that this enzyme consumes. Reduction of one single molecule of dinitrogen by the nitrogenase complex uses 16 ATP molecules or 2.66 per electron transferred to a nitrogen atom. However, with the consumption of reducing power to get the eight electrons, 25 ATP equivalent are consumed during one single reaction. This reaction is therefore one of the most energy consuming of the living (Simpson & Burris, 1984). As shown in the above reaction, hydrogen (H₂) is produced as a by-product of the N₂ reduction, adding some wasted energy and reducing power cost to the overall process. Some bacteria possess *hup* (hydrogen uptake hydrogenase) genes, which help

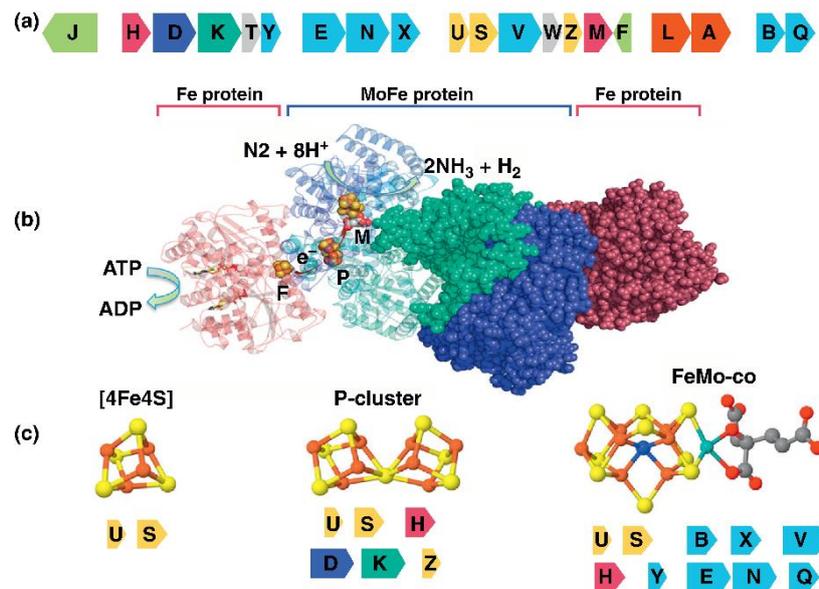


Figure 15 | Genomic structural and functional organization of the nitrogenase. This figure depicts the gene organization of the *nif* (nitrogen fixation) cluster from *Klebsiella oxytoca*. **a** | Genes necessary to the synthesis of the whole nitrogenase complex are listed here. The genes are colored after the crystal structure of the corresponding proteins (NifHDK) or their function (all other genes). **b** | Structure of the nitrogenase complex. Half of the complex is presented in transparent ribbons, revealing the position of every cofactor. **c** | Structure of the nitrogenase cofactors. Genes involved in their synthesis are mentioned. Extracted from Oldroyd & Dixon, 2014.

them to recycle H_2 , thus improving energy performance of nitrogen fixation (Fernández et al., 2005; Masuda et al., 2016).

Among all the nitrogenase complexes, the best studied is the one of the bacteria *Klebsiella oxytoca*, whose structure and assembly are well described. This enzymatic complex is composed of six subunits encoded by three genes but the whole complex can be split into two metalloproteins. One is the dinitrogenase reductase or Fe-protein, a homodimer encoded by the *nifH* gene. The other one, the dinitrogenase $2\alpha 2\beta$ or Mo-Fe protein is a heterotetramer encoded by the genes *nifD* for the α subunit and *nifK* for the β subunit (**Fig. 15a&b**) (Peters et al., 1997). Each of the metalloproteins is bound to iron-sulfur clusters involved in the electron transfer activity of this enzyme (**Fig. 15c**). They convey electrons resulting from the ATPase activity of the Fe-protein towards the Mo-Fe catalytic center. These iron-sulfur clusters are synthesized by the NifUS(Z) proteins (Seefeldt et al., 2009). The last electron acceptor of the complex is a Fe-Mo-Co cofactor, short for [7FeMo9s:C]-homocitrate (**Fig. 15c**) (Lancaster et al., 2011). Molybdenum is brought thanks to the NifQ protein and homocitrate is synthesized by NifV, the homocitrate synthase before being passed to the nitrogenase by NafY (Rubio &

Ludden, 2008). However, *nifV* is not an ubiquitous gene among bacteria that are able to fix nitrogen. In fact, it is present only in free-living nitrogen-fixing bacteria (true diazotrophs). These can synthesize the whole Fe-Mo-Co cofactor and be autonomous for nitrogen fixation. Yet, most of the plant symbionts do not possess this gene and thus are bound to their host for nitrogen fixation. In this scenario, the plant provides homocitrate to its microsymbiont, so that the nitrogenase can be functional. In *L. japonicus*, the *FEN1* gene encodes the enzyme responsible for the synthesis of homocitrate from Acetyl-CoA and Alpha-ketoglutarate (Hakoyama et al., 2009). It is even possible for a strain that possesses *nifV* that is artificially mutated to perform a functional symbiosis with its host if this one has a copy of *FEN1*. The study of Nouwen et al. 2017 describes this particularity with the *Bradyrhizobium* strain ORS285 in interaction with *A. afraspera*. The only rhizobia that possess the *nifV* gene are some photosynthetic *Bradyrhizobium* strains and *A. caulinodans* (Alazard, 1990; Dreyfus et al., 1988; Nouwen et al., 2017). *nifV* is not the only gene that is absent in some rhizobia, as for instance some *Sinorhizobium* species only count 8 of the *nif* genes. However, this variation in the number of *nif* genes has been supposed to be linked to the difference in bacterial physiology. Under this hypothesis, unknown functions could complement the lack of these genes (Masson-Boivin et al., 2009; Rubio & Ludden, 2008).

Nitrogenase sensitivity to dioxygen – How is it overcome?

One feature of the nitrogenase complex is its sensitivity to dioxygen. When the O₂ concentration in the nodule is too high, the complex becomes irreversibly inactive because of the oxidation of the iron-sulfur clusters. However and paradoxically, as usually being aerobic organisms, rhizobia require oxygen in order to perform the cellular respiration and more precisely for the oxidative phosphorylation that is required for energy production. Nitrogen fixation by rhizobia can neither take place in aerobic nor anoxic conditions. Therefore, O₂ tension in nodule must be comprised between 5 and 30 nM to have functional nitrogenases and ATP-producing bacteria, a condition that is called microoxy (Marchal & Vanderleyden, 2000; Terpolilli et al., 2012). To overcome this challenge, physical and chemical barriers are set up in nodules. The first barrier is physical as nodule cortex and endodermis act as a first filter, partially preventing the diffusion of O₂ (Minchin et al., 2008). The second barrier is of chemical nature and is due to various proteins produced by both symbiotic partners.

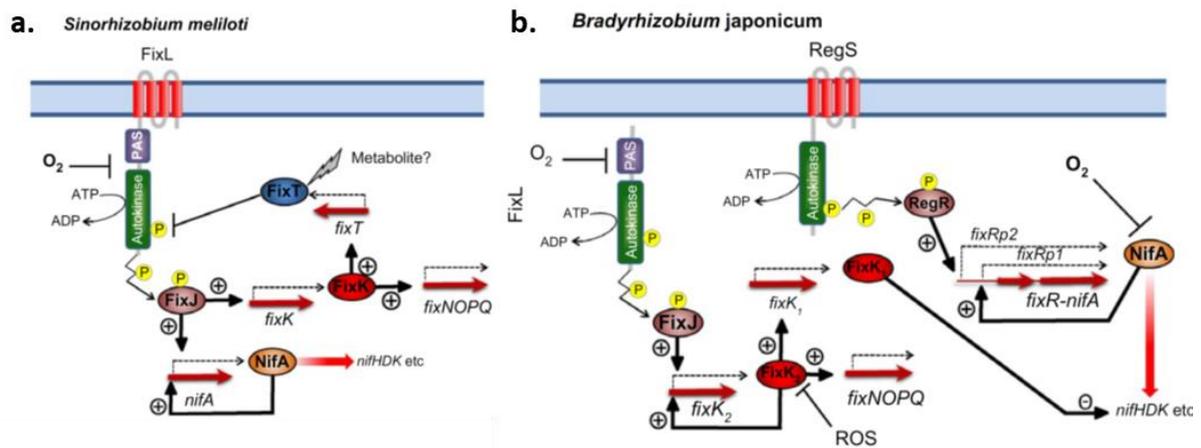


Figure 16 | Regulation of bacterial symbiotic gene expression by O_2 tension. a | Action of FixLJ in *Sinorhizobium meliloti*. When oxygen tension is low enough, the two-component system FixLJ will activate the expression of the *fix* operon as well as *nifA*. This factor is itself sensitive to O_2 and activates the expression of the *nif* genes. **b |** FixLJ system works differently in *Bradyrhizobium diazoefficiens* USDA110. Most notably is the activation of *nifA*, which is independent of FixLJ regulation. In this bacterium, *nifA* is under control of RegSR. Extracted from Terpolilli et al., 2014.

On one side, the plant produces leghemoglobin, an iron heme protein that acts similarly to human hemoglobin, having a high affinity to oxygen and chelating it (Udvardi & Poole, 2013). The concentration of free dioxygen is reduced thus preventing it to react with iron-sulfur clusters. The leghemoglobin is the molecule that gives nodules their pinkish color, almost red when cut open. This color is usually an indicator of a functional symbiosis as the absence of leghemoglobin is a hallmark of nitrogen fixation failure.

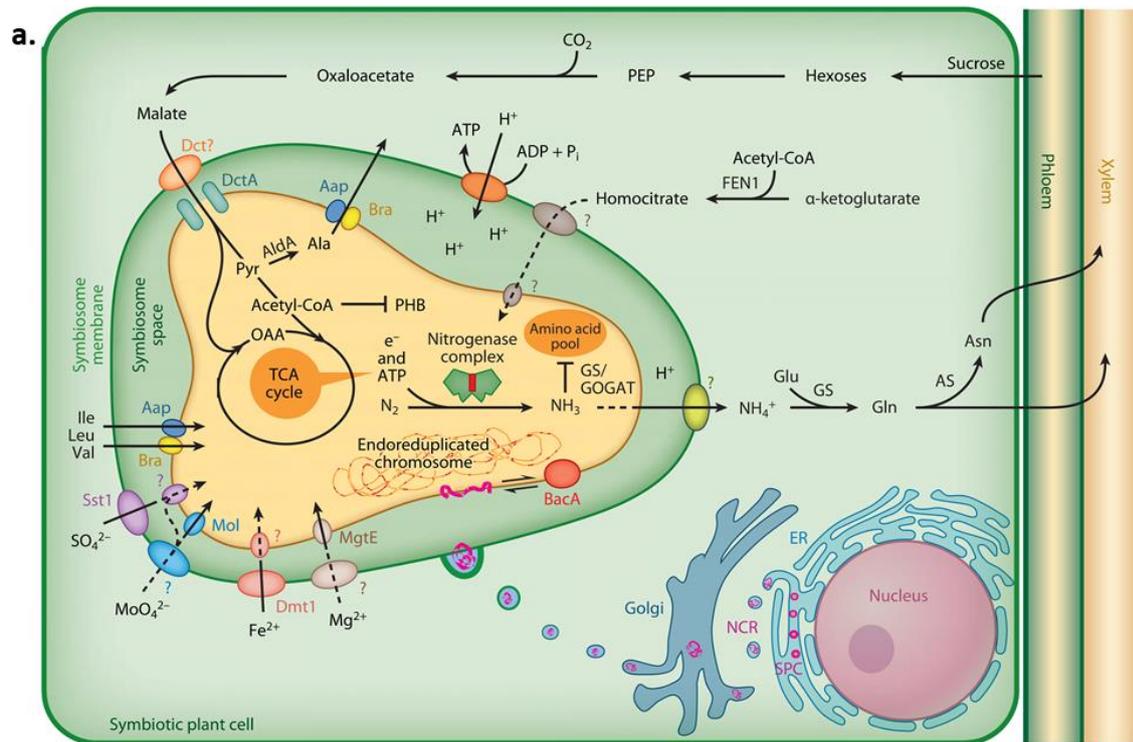
On the other side, the *fix* genes play an important role in the behavior of bacteroids depending on O_2 levels. First, the two-component system FixLJ is sensing dioxygen levels (**Fig. 16**). The protein FixL is a histidine kinase with a heme-based oxygen sensor. In absence of oxygen, the protein will phosphorylate the FixJ response regulator. After dimerization, FixJ will induce the expression of several genes, but those, which will hold our interest here, are *fixK* and *nifA*, which encode transcriptional regulators. FixK will activate the expression of the *fixNOPQ* genes (**Fig. 16**) (Terpolilli et al., 2012). This set of genes encodes for a cytochrome cbb3-type oxidase which reduces dioxygen to either H_2O or hydrogen peroxide thanks to its oxygenase activity (Delgado et al., 1998; Thöny-Meyer, 1997). The oxidase is part of a respiratory chain, which is specific for life under microoxic conditions. Hence, their affinity for O_2 is far stronger than the usual respiratory chain that is active in free-living rhizobia. Reducing dioxygen both allow to regulate O_2 tension within the bacteria and produces reducing power

for the nitrogenase activity. NifA, which is itself regulated by O₂ tension, activates the expression of the *nif* genes, among which *nifHDK* encode the nitrogenase complex. Thus, the nitrogenase is expressed only when O₂ tension is low enough for the nitrogenase to be active. Other *fix* genes are important such as *fixABCX* or *fixGHIS* (Terpolilli et al., 2012). The first set of genes is required for a functional symbiosis; however, their roles are not fully understood yet. They might be important for conveying electrons towards the nitrogenase. The second set of genes, *fixGHIS* are important for the assembly of the FixNOPQ oxygenase and are thus essential during symbiosis.

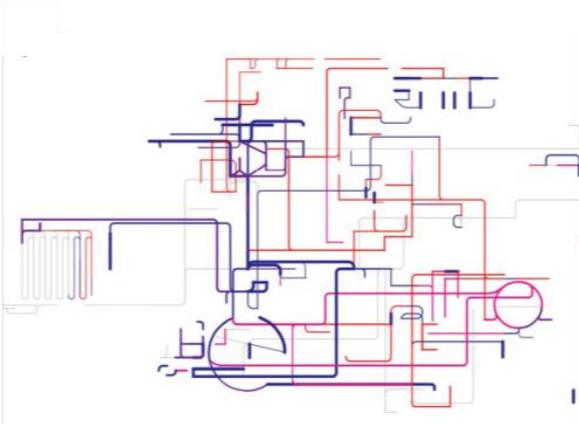
A nutritional symbiosis is a matter of metabolite exchanges

As described earlier in this manuscript, plants trade carbon sources for fixed nitrogen during the mutualistic symbiosis with rhizobia. Thanks to nitrogenase, bacteroids produce NH₃ from N₂ and export it to the plant cells. The way ammonia is transported to plant cells is not known yet but it likely enters the plant cytoplasm under the form of ammonium (NH₄⁺) (**Fig. 17a**). This form of nitrogen is then assimilated by the glutamine synthase (GS), which binds it with glutamate, producing glutamine. This pool of glutamine can be transformed into asparagine by the asparagine synthase (AS). Then, asparagine or glutamine will transit through the vascular tissues of the xylem to be used by the rest of the plant (Oldroyd et al., 2011; Udvardi & Poole, 2013). In turn, the plant provides bacteroids with carbon sources under the form of dicarboxylates, often malate or succinate that will fuel the bacterial tricarboxylic acid (TCA) cycle for cellular energy production. While the transport system is yet unknown on the plant side, the dicarboxylate transporter DctA in the bacterial envelope is taking up dicarboxylates (**Fig. 17a**).

Restraining the exchanges between plant and bacteroids to only these molecules would be missing some of the fluxes that underpin the symbiotic interaction. Indeed, though the metabolism of bacteria is highly specialized compared to bacteria living in bulk soil, they need additional nutrients to survive and perform their function (DiCenzo et al., 2016). As described before, in most of the cases, plants must provide their symbionts with homocitrate in order to have a functional nitrogenase. Besides this function-driven requirement, bacteria receive ions from the plant, such as K⁺ and Ca²⁺ (Roberts & Tyerman, 2002), Fe²⁺ (Kaiser et al., 2003), Cl⁻



b. Rhizosphere bacteria metabolism



c. Bacteroid metabolism

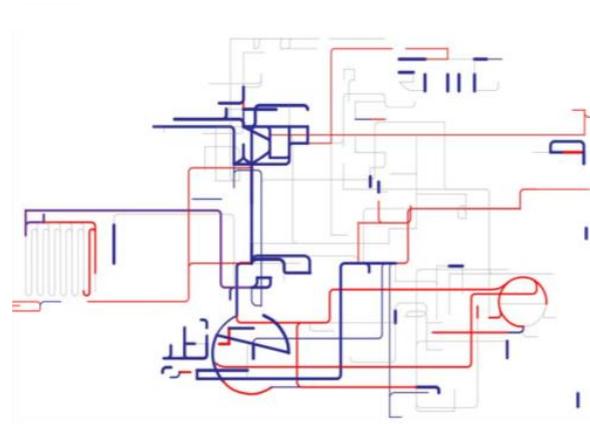


Figure 17 | Bacteroid metabolism and exchanges with the plant. a | Schematic of the metabolite exchanges happening in a symbiotic cell between the host plant and bacteroids. Besides dicarboxylates, the plant provides the bacterium with several minerals and amino acids to fulfill its metabolic needs, in order to maximize nitrogen fixation and export to the plant. **b & c** | Metabolic fluxes of *Sinorhizobium meliloti* in rhizosphere (**b**) or *in planta* (**c**) lifestyles generated from the metabolic model of this bacterium. Lines represent metabolic pathways, their thickness the flux through each reaction and color the essentiality of the reactions. The thicker the line greater is the flux (in log scale). Grey lines indicate unused pathways. Blue lines indicate a fitness decrease of less than 1% when reactions are lost; dark purple a decrease of less than 50%; bright purple a decrease of more than 50%; and red a fitness decrease of more than 99%. Bacteroid metabolism is more specialized than rhizosphere bacteria. Extracted from Oldroyd et al., 2011 (**a**) and DiCenzo et al., 2016 (**b & c**).

(Szczyglowski et al., 1998; Vincill et al., 2005), SO_4^{2-} (Krusell et al., 2005), Zn^{2-} (Moreau et al., 2002), MoO_4^{2-} , Mg^{2+} and H^+ (Oldroyd et al., 2011), required for the nitrogenase activity or housekeeping functions (**Fig. 17a**). Various transporters allowing their import into bacteroids were described though some of them are not identified yet as shown in **figure 17a**. Last but not least, bacteroids also need some of the elementary bricks of the living: amino acids. Indeed and strikingly, symbionts become unable to synthesize the branched amino-acids (isoleucine, leucine, valine) whereas they are still required for a functional symbiosis (Prell et al., 2009; Udvardi & Poole, 2013). This phenomenon, called symbiotic auxotrophy is explained by the high level of metabolic specialization bacteria undergo when they become bacteroids. In fact, several functions that are required in the soil are unused during symbiosis. **Figures 17 b & c** displays the behavior of *S. meliloti* 1021 metabolic model in two growth conditions: rhizosphere (**b**), or nodule (**c**). In the nodule, metabolite fluxes are focalized on few pathways, maximizing the production of ammonia. The bacteroid metabolism is hyperspecialized, and this feature very well depicts the concept of symbiotic auxotrophy. Indeed, during symbiosis the bacteroid-plant cell systems function in tense and continuous fluxes, trying to find an equilibrium state between the two partners to minimize loss of energy for the plant. Ideally, every molecule given by the plant should be used by bacteroids. However, in some cases the fluxes that transit between the two partners somehow desynchronize, often in the case of a malfunctioning symbiosis. Accumulation of various compounds can be observed in these circumstances such as starch granules in plant cells that are usually considered as stress markers. Another marker of a malfunctioning symbiosis is the accumulation of polyhydroxybutyrate (PHB) in bacteroids for similar reasons than starch granules. Indeed, both starch and PHB are storage compounds, which are usually not seen in fully functional and unstressed nodules.

The lasts stages of the interaction

A senescence zone appears a few weeks after nodule formation. In this zone, bacteroids do not fix nitrogen anymore and are progressively digested from center to periphery by the plant (Dupont et al., 2012; Van de Velde et al., 2006). It may be possible that dead bacteroids become a source of nutrients for the plant. From there, leghemoglobin is degraded, releasing biliverdin, a by-product of this degradation, which gives the senescence zone a greenish color. It is

thought that this degradation not only promotes senescence through inactivation of the nitrogenase it could also induce the formation of Reactive Oxygen Species (ROS) which would help the process of senescence (Dupont et al., 2012). Plants cells eventually begin to be digested. It has been shown that proteases such as caspases or Cysteine proteases are expressed in the senescence zone. They would help the degradation of both host and symbiont cells (Pierre et al., 2014; Van De Velde et al., 2006). In even older nodules, a fifth zone can appear. This zone is the saprophytic zone where bacteria released by infection threads may degrade and feed on the dying plant cells. Cell-death markers begin to appear such as plasma membranes losing their integrity and digestion of cell walls (Van De Velde et al., 2006). These bacteria called "rhizoboids" are saprophytic and may feed from plant cells remains (Timmers et al., 2000). The formation of these zones keeps going until the whole nodule becomes a mix of senescence and saprophytic zones. Nodule senescence occurs naturally when nodules age, but it can also be induced by environmental factors such as dark stress or adjunction of nitrogen in the growth medium.

D. The Bacterial differentiation within nodules

In some legume lineages, the release of bacteria into the plant cell cytoplasm, is followed by an extreme bacterial differentiation process known as terminal bacteroid differentiation (TBD) that is explained in the following section. This happens for example in plants from the above described IRLC and Dalbergioid legumes. Terminally differentiated bacteroids are also called swollen bacteroids. Oppositely, in other legume species, bacteroids remain morphologically similar to free-living bacteria. As such, this type of bacteroids is called unmodified (U-type) or non-swollen. Bacteroids of *M. loti* and *B. diazoefficiens* USDA110 in nodules of *L. japonicus* and *G. max*, respectively, are examples of U-type bacteroids (Oono & Denison, 2010).

i. Bacteroid features are modified to reach the TBD

During TBD, the most visible change in the bacteroids is their morphology (**Fig. 18a**). As opposed to U-type bacteroids, terminally differentiated symbionts undergo a cell volume increase and become either elongated (E-type) or spherical (S-type). IRLC plants usually host E-type bacteroids, such as in *Vicia sativa* (Beijerinck, 1888) or *M. sativa* (Vasse et al., 1990) for example. E-type bacteroids keep their rod shape but their length can increase up to 10 times their initial length (0.5-1 μm to 10 μm) (Mergaert et al., 2006; Montiel et al., 2017; Van de Velde et al., 2010). This elongation usually results from polar growth, sometimes with branching, resulting in Y-shaped bacteroids (Mergaert, 2018). In *Dalbergioids*, depending on the plant, bacteroids are either of the E-type or the S-type. *Arachis hypogaea* and *Aeschynomene indica* are examples of plants inducing the S-type bacteroid differentiation (**Fig. 18c**) (Bonaldi et al., 2011; Oono & Denison, 2010). S-type bacteroid ensue from a polar growth followed by a loss of the initial morphology thus reaching the spherical shape (Czernic et al., 2015). Bacteroids can reach a 2-3 μm diameter in *A. indica* at 13 dpi (Bonaldi et al., 2011), or in *Ononis spinosa*, one of the few known IRLC hosting S-type bacteroids (Montiel et al., 2016). In indeterminate nodules, it is possible to keep track of the process of differentiation as a progressive elongation of the bacteroid can be observed from the infection zone (II) through the end of the fixing zone (III) (Vasse et al., 1990). This change in morphology implies a induction of membrane synthesis genes expression and a modification of the envelope composition (Mergaert, 2018).

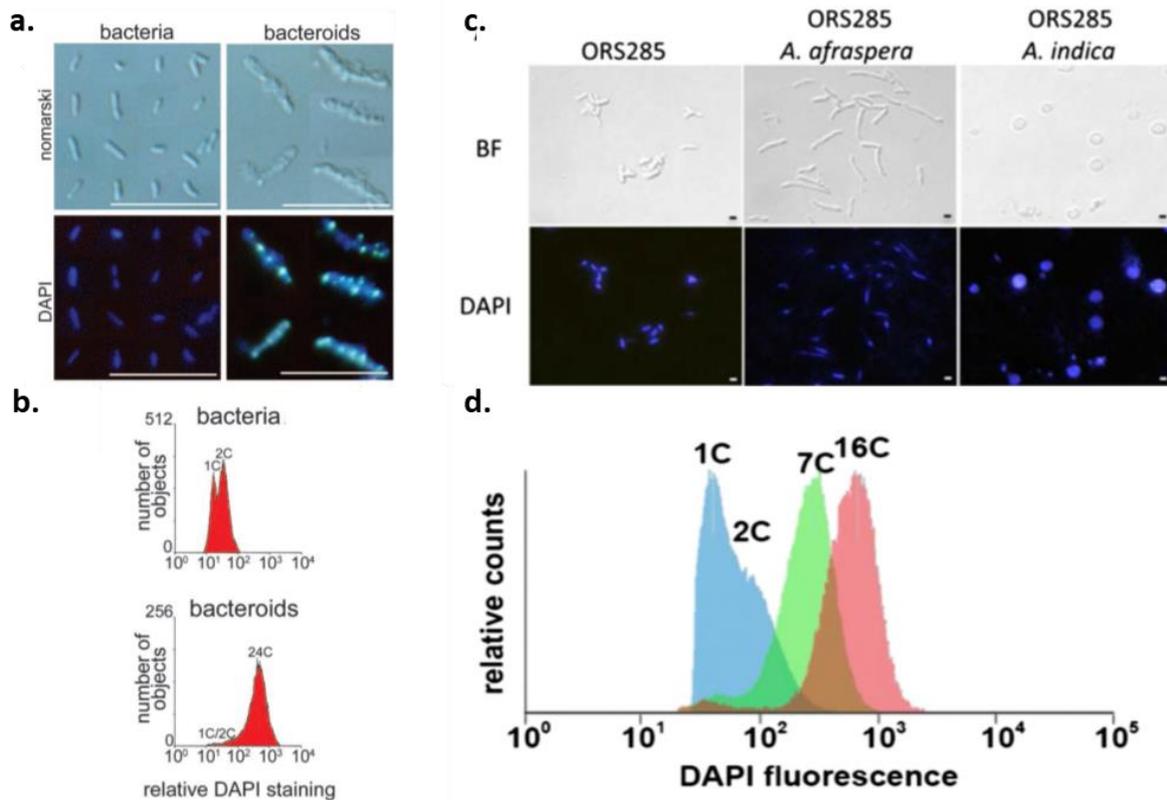


Figure 18 | Terminal bacteroid differentiation features. Assessment of *Sinorhizobium meliloti* TBD features in interaction with *Medicago truncatula* (a & b) or of *Bradyrhizobium* sp. ORS285 in interaction with either *A. afraspera* or *A. indica* (c & d). In all cases, bacteroids undergo endoreduplication, have an increased cell volume and become either elongated or spherical. a & c | Morphology of enlarged bacteroids extracted from mature nodules compared to culture bacteria. Note that endoreduplicated *S. meliloti* bacteroids display numerous nucleoids. Scales bars, a: 10 μm ; c: 1 μm . b & d | Ploidy level of free-living bacteria versus endoreduplicated bacteroids measured by flow cytometry after DAPI staining. Extracted from Mergaert et al., 2006 (a & b) and Czernic et al., 2015 (c & d). d | Colors refers to the strain ORS285 either in culture (blue), or purified bacteroids from *A. afraspera* (green) or *A. indica* (red).

Cell elongation is not the only change that occur during TBD. Membranes of swollen bacteroids has an increased permeability due to the action of a massive amount of plant-produced antimicrobial peptides (AMP) (Haag et al., 2011; Mikuláss et al., 2016). Another change is the increase of the ploidy level of bacteroids. Indeed, during their differentiation, the cell cycle of bacteroids is altered. Instead of normally dividing after DNA replication, bacteria skip the cytokinesis phase but keep replicating their genome, thereby entering endoreduplication cycles. Their ploidy level can reach up to 24 DNA copies per cell in *Medicago* (Fig. 18b) (Mergaert et al., 2006) or 16C in *A. indica* (Fig. 18d) (Czernic et al., 2015). This phenomenon suggest that bacterial cell cycle regulation is not functioning normally during

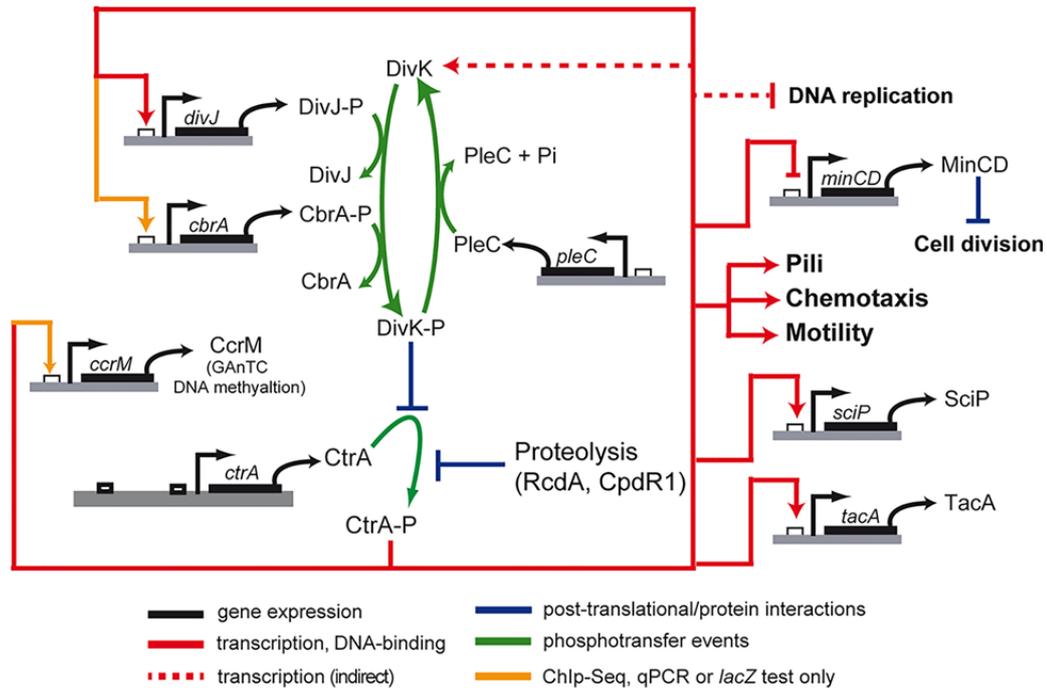


Figure 19 | Cell cycle regulation in *Sinorhizobium meliloti*. This scheme represents the regulation network of the cell cycle by CtrA in *Sinorhizobium meliloti*. The color code explained in the legend gives information on which type of data was used to generate this scheme. Note that CtrA is regulated at the transcriptional level, and at the post-translational level by both phosphorylation and proteolytic degradation. As CtrA promotes cell division and represses genome replication, its level is dropping down *in planta* in differentiating bacteroids that undergo endoreduplication. Extracted from Pini et al., 2015.

TBD. CtrA is known as the master regulator of the cell cycle in Alphaproteobacteria and has been

well described in *Caulobacter crescentus*, a model for the study of cell cycle, and *S. meliloti* (Biondi et al., 2006; Pini et al., 2015; Xue & Biondi, 2019). The role of this response regulator, which is active when phosphorylated, is to promote cytokinesis while inhibiting DNA replication (**Fig. 19**). This protein is controlled by several levels of regulation, transcriptionally and post-translationally through phosphorylation and degradation. Therefore, its expression and activity usually vary during the cell-cycle, from very low levels during replication to high levels during cytokinesis. This precise dynamic variation of CtrA activity allows tight coordination and the correct sequence of cell-cycle phases. In *M. truncatula* bacteroids, *ctrA* expression gradually decrease from the infection zone (II) to the fixing zone (III) (Roux et al., 2014) and CtrA protein levels drop strongly in bacteroids (Pini et al., 2015). Knock-out mutation of *ctrA* is lethal for the bacteria, but artificial depletion of CtrA leads to a bacteroid-like morphology (Pini et al., 2015).

Without CtrA, repression of replication is lifted, and promotion of cytokinesis abolished. These effects corroborate with the polyploid bacteroid phenotype. In *S. meliloti*, other cell cycle mutants show similar phenotypes. It is the case of the *divJ* and *cbrA* mutants, which are both indirect repressors of CtrA activity. The same effect is obtained in PleC over-expressor strains, PleC being an indirect inducer of CtrA activity (**Fig. 19**). The endoreduplication of swollen bacteroids is also responsible for the loss of the bacterial ability to return to free-living growth as they cannot manage a division with multiple copies of their chromosomes. When extracted from nodules and plated on a rich medium, swollen bacteroids cannot grow. Therefore, this differentiation is called terminal because of its irreversible nature.

ii. NCR peptides as orchestrators of this symbiotic process

TBD is not a universal process in the legume-rhizobium symbiosis. As an example, *Rhizobium leguminosarum* can differentiate both to swollen or unswollen bacteroids depending on its host plant. If it interacts with *V. sativa* or *P. sativum* that belong to the IRLC, it will end up as swollen bacteroids. Conversely, if the host plant is *P. vulgaris*, it will form unswollen bacteroids (Mergaert et al., 2006). Indeed, only plants from some legume clades, including the IRLC and *Dalbergioid* clades, host swollen bacteroids. These plants are the only ones able to synthesize a specific family of AMP called NCR peptides (Nodule-specific Cysteine-Rich) in IRLC or NCR-like in *Dalbergioids*. Besides their antimicrobial activity, these short peptides (60-90 amino acids or 30-50 after cleavage of their signal peptide), have been shown to be responsible for triggering TBD (Van de Velde et al., 2010). They are specifically expressed in nodules (**Fig. 20b**), contain a motif composed of cysteine residues in fixed positions which number can vary depending on the plant (**Fig. 20a**), a conserved signal peptide, but except these features, the overall sequence is rather divergent (Mergaert, 2018).

NCR peptides in IRLC

NCR peptides were first discovered in IRLC legumes. *M. truncatula* possesses a large diversity of peptides as more than 600 different NCR peptides were identified in the genome of this species (Mergaert et al., 2003; Montiel et al., 2017). Proteomic analysis identified 138 NCR peptides in *Sinorhizobium* bacteroids. Among them, 38 are cationic (pI: 7.51-9.6), but the majority is either anionic (pI: 3.4-6.5) or neutral (pI: 6.51-7.5) (Durgo et al., 2015). The

localization of the NCR peptides in the bacteroids was further confirmed by immunolocalizations or the expression of NCR fusions with fluorescent proteins (Van de Velde et al., 2010; Kim et al., 2015; Horvath et al., 2015; Yang et al., 2017; Wang et al., 2017) and demonstrates that these plant produced peptides target the bacteroids. All analyzed species of the IRLC produce NCR peptides (Montiel et al., 2016, 2017). Interestingly, there is a positive correlation between NCR number and bacteroid cell length (Montiel et al., 2017). Additionally, plants that go not induce terminally differentiated bacteroids, such as the above-mentioned *L. japonicus*, *G. max* or *P. vulgaris* do not have NCR genes (Mergaert et al., 2003). Together, these correlations suggested a role of the NCRs in the TBD process. This hypothesis was experimentally confirmed by different observations. The prevention of the transport of NCRs to the bacteroids leads to no TBD, while on the contrary, expressing *NCR* genes in *L. japonicus*

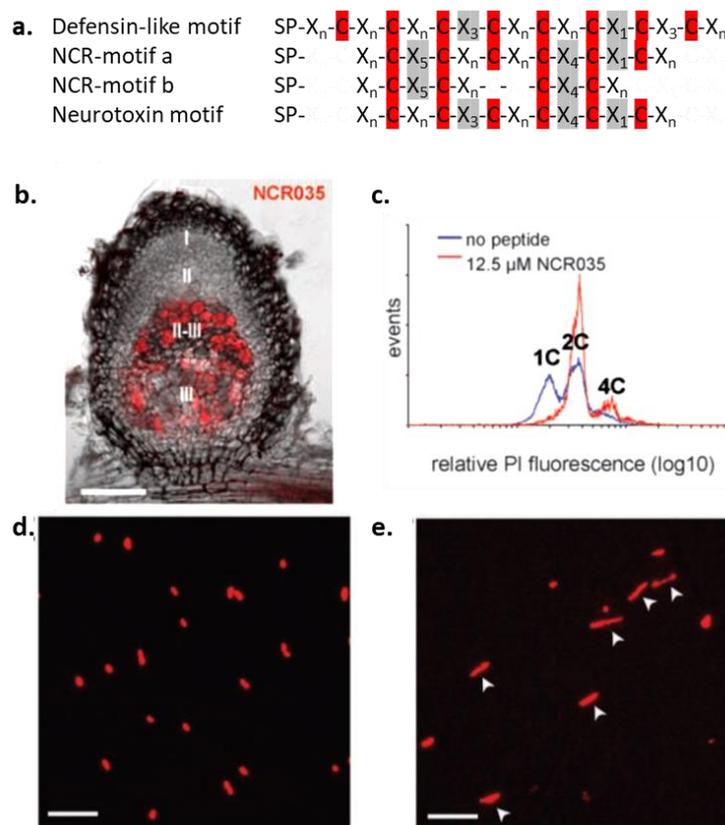


Figure 20 | NCR peptides and their TBD-inducing activity. a. | NCR motifs compared to defensin and neurotoxin motifs. **b.** | Example of NCR peptide localization. NCR035 was detected by immunofluorescence in the inter-zone and the nitrogen-fixing zone. **c.** | Effects of NCR035 peptide treatment on the ploidy level of *Sinorhizobium meliloti* cultures. DNA content of the cells was assessed by measuring propidium iodide (PI) fluorescence levels with flow cytometry. Treatments with synthetic peptides is sufficient to induce polyploidization. Scale bar: 100 μm. **d & e** | Effects of NCR035 peptide treatment on the morphology of the bacteria. Bacterial cultures in control conditions (**d**) and elongated bacteria treated with NCR035 (**e**). Scale bar: 10 μm. **b, c, d & e** extracted from Van de Velde et al., 2010.

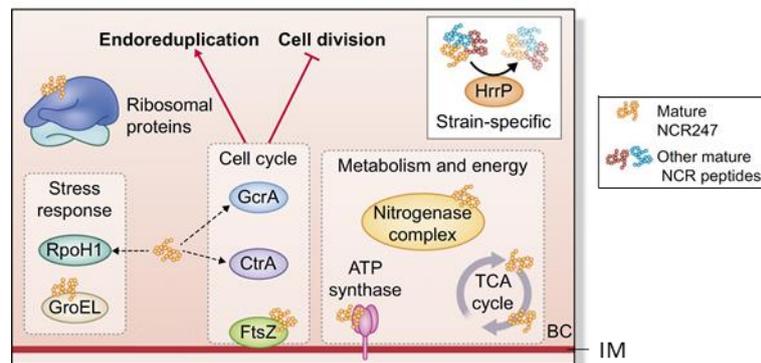


Figure 21 | Intracellular targets of NCR peptides. This scheme describes the targets of NCR peptides in the cytosol of bacteroids, as well as bacterial functions which interact with NCR peptides. Indeed, stress response and primary metabolism enzymes were shown to interact directly with NCR peptides. Cell-cycle regulation enzymes were shown to be regulated by NCR treatments though direct interaction has not been demonstrated yet. All those functions are thought to be targets of NCR peptides, but on the other side, some bacteria (HH103) developed functions which target NCR peptides, such as the HrrP protease which digests NCR peptides. IM: inner membrane; BC: bacteroid cytoplasm. Adapted from Alunni & Gourion, 2016.

nodules leads to bacteroids with TBD features, it is also possible to mimic TBD with treatment of bacterial cultures with certain NCRs. In this case, bacteria undergo both cell elongation and endoreduplication, although to a limited extent as compared to bacteroids (**Fig. 20 c, d & e**).

Most of NCR peptides are cationic AMPs and as such they can disrupt the integrity of bacterial envelope but also the membranes of some fungi, leading to a loss of membrane potential (Mikuláss et al., 2016; Tiricz et al., 2013). However, killing the symbionts is not what occurs during symbiosis. How NCRs control the TBD is not fully understood yet, but some of the effects of NCRs on the bacterial physiology are known though. Treatment of *S. meliloti* cultures with NCR247 induces various changes in the bacterial life and modulate the expression of a large set of genes (Penterman et al., 2014; Shabab et al., 2016). For example, treatment of culture bacteria with this NCR is enough to reduce the expression of *ctrA* to very low levels, but also of other cell cycle regulators, thereby delaying the cell cycle progression (**Fig. 21**) (Penterman et al., 2014; Shabab et al., 2016). A direct interaction of NCR peptides with CtrA was not shown until today though. However, it is likely NCRs somehow alter the cell cycle, either through CtrA directly or through one of its regulators. Moreover, NCR247 was shown to inhibit protein synthesis and interacts with different proteins in bacteria. Besides primary metabolism enzymes, one of them is FtsZ a protein involved in the cell division (**Fig. 21**). In *S. meliloti*, two copies of FtsZ coexist: FtsZ1, which is essential for cell survival, and FtsZ2, a

truncated and dispensable version of FtsZ1. FtsZ1 can polymerize into a contractile circular structure called Z-ring, localized at the site of cytokinesis. This structure is required for the cell division as it helps recruiting other proteins that form a complex called divisome. While the divisome builds the septum, the Z-rings contracts to separate the two daughter cells (Erickson et al., 2010). Thus, NCRs could possibly block cell division by acting directly on divisome

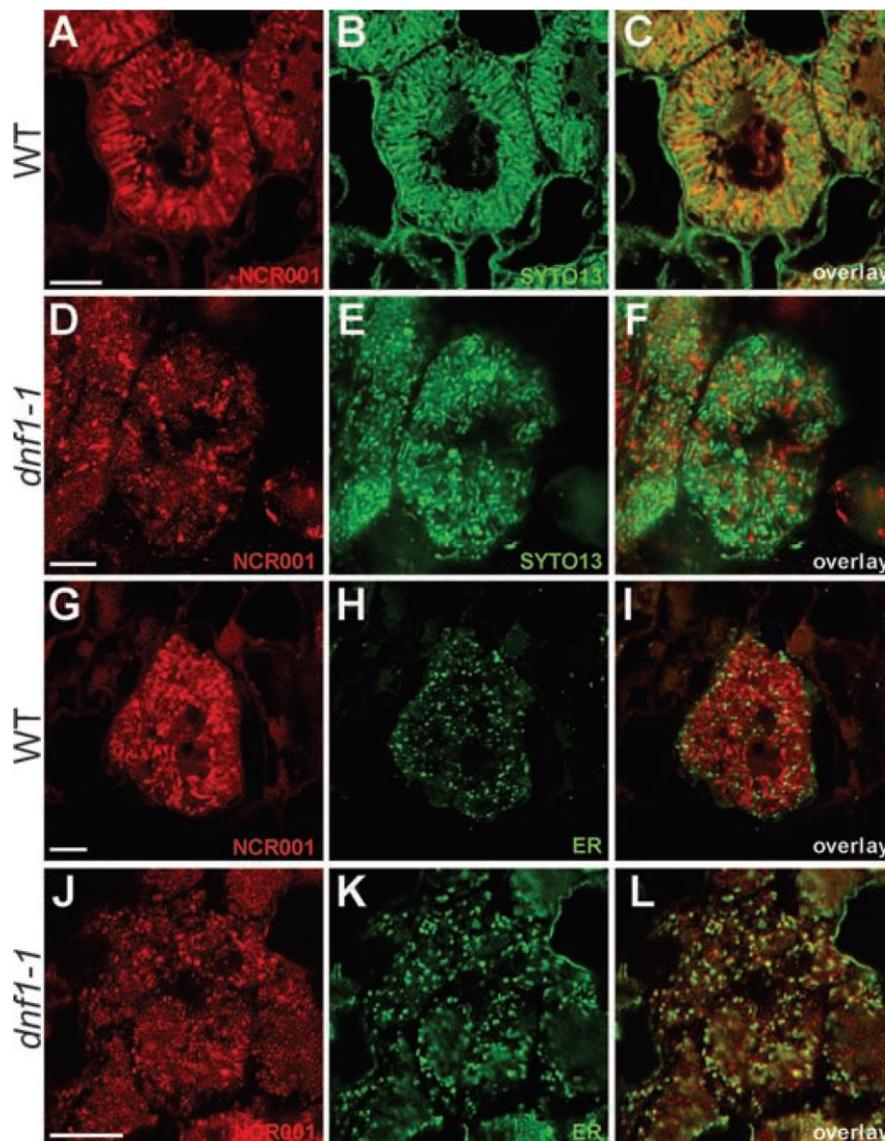


Figure 22 | Effects of the mutation of the signal peptidase DNF1 on bacteroid differentiation. Three different staining were used in the presented study. Immunofluorescence staining of NCR001 is shown in red (**A, D, G & J**). In green, either SYTO9 a nucleic acid marker which here stains bacteria (**B & E**) or immunofluorescence targeting ER (**H & K**). NCR001 was used to assess the effects of the *dnf1-1* mutant. In WT plants, NCR001 colocalizes with elongated bacteroids (**A, B & C**). However, mutation of DNF1 leads to the loss of the superposition of NCR001 signal with SYTO9 signal and bacteroids remain undifferentiated (**D, E & F**). Indeed NCR001, instead of being targeted to symbiosomes is sequestered in the ER (**J, K & L**). Usually, no colocalization of NCR001 and ER is observed (**G, H & I**). Scale bars: 10 μm . Extracted from Van de Velde et al., 2010.

assembly and/or functioning. In bacteroids, NCR247 was also found to interact with the chaperonin GroEL (**Fig. 21**), which is essential for a functional symbiosis, with some subunits of the nitrogenase but also with other NCR peptides (NCR028, NCR169 & NCR290) (Farkas et al., 2014). This last finding suggest some NCRs might form heterocomplexes.

NCR peptides are expressed by the plant in nodules and they target symbiosomes to perform their antimicrobial/TBD-inducing activity. To do so, they are produced and translocated using the classical secretion machinery of eukaryotic cells (Mergaert et al., 2003). NCR peptides are translated by ribosomes, which are docked onto the endoplasmic reticulum (ER). The peptides are thus directly synthesized into the ER where they fold to get their 3D structure. They are then transferred to the Golgi apparatus after cleavage of their signal peptide by the signal peptidase complex comprising DNF1. From the Golgi, NCR peptides are internalized into secretory vesicles. After being transported close to symbiosomes thanks to the cytoskeleton, the vesicles eventually fuse with the symbiosome membrane thus releasing NCRs into the peri-bacteroid space (Alunni & Gourion, 2016; Van de Velde et al., 2010). Interestingly, a genetic screening of a *M. truncatula* mutant library led to the discovery of three NCR-defective mutants. These mutants are part of seven *fix⁻* mutants named *dnf* (defective in nitrogen fixation) after their *in planta* phenotype (Starker et al., 2006). *dnf1* is the mutant with the strongest phenotype. Indeed, the mutation affects a Signal Peptidase Complex (SPC) subunit, expressed in the nodule, which is responsible for the cleavage of signal peptides, including those of the NCRs (Wang et al., 2010). Therefore, when *DNF1* is mutated no peptide can be secreted through the classical secretion machinery in nodule cells. NCR peptides, which usually co-localize with symbiosomes, are not matured and remain sequestered into the ER in the context of the *DNF1* mutant and no TBD occurs (**Fig. 22**) (Van de Velde et al., 2010).

The two other *dnf* mutants, which are linked to NCR peptides, are *dnf4* and *dnf7*. The mutated genes respectively encode for NCR211 and NCR169 (Horvath et al., 2015; Kim et al., 2015). Both mutants have impaired symbiosis characterized by an early senescence of the symbiotic organs along with bacteroid death occurring either before (in *dnf4*) or after (in *dnf7*) TBD. It is striking that two single NCR lead to such strong phenotypes when mutated. Indeed, the large array of NCR peptides in *M. truncatula* could have suggested that there is redundancy among them. While this may be the case for some NCRs, it appears some other NCRs are solely

responsible for a specific function. This feature might reflect the complexity of the TBD process, as well a diversity of functions of NCRs. Some NCR peptides are also responsible for selecting symbionts, restraining the range of compatible strains. NFS1 and NFS2 NCR peptides have this role in *M. truncatula*. The A17 ecotype possesses *NFS1*⁻ and *NFS2*⁻ alleles and is homozygous for both *NFS1* and *NFS2* loci (Wang et al., 2017; Yang et al., 2017). In this genotype context, the *S. meliloti* Rm41 strain cannot maintain a functional nitrogen-fixing zone as infected plant cells quickly undergo lysis after bacterial differentiation while other *S. meliloti* strains are not affected by these two NCR alleles. Oppositely, the DZA315 *M. truncatula* ecotype possess homozygous copies of *NFS1*⁺ and *NFS2*⁺ alleles. In this context, Rm41 can carry out a fully functional symbiosis with its host. These findings raise the following question: how many functions to control symbiosis do NCRs cover? The high diversity of NCR peptides suggest that other functions are yet to be unveiled. Other NCR peptides have been studied, mostly cationic ones. Indeed, this class of NCR peptides have the strongest activity, for example as antimicrobial agents. The activity of neutral and anionic NCR peptides is badly known today. However, NCR peptides are sequentially expressed in nodules and sometimes specific to different nodules zones (**Fig. 23a&b**). It is possible that cationic NCRs are responsible for permeabilizing the symbiont's envelope in order for other peptides (anionic and neutral) to enter the bacterial cells and carry out different activities within bacteroids.

In *M. truncatula*, two different cysteine motifs were identified in the sequence of NCR peptides with either four or six cysteine residues in fixed positions. These motifs resemble defensin or neurotoxin patterns (Mergaert et al., 2003). The conserved cysteines are largely involved in the activity of NCRs and notably by contributing to their electrical charge. Likewise, their redox state appears to be important for their activity (Haag et al., 2012; Shabab et al., 2016). While in the ER, NCRs are likely to be oxidized before being secreted into symbiosomes. There, NCRs are co-secreted with thioredoxins. Thioredoxins are a family of proteins involved in the reduction of protein disulfide bonds, among other molecules, thus freeing bound cysteines. It has been shown that thioredoxin s1 uses NCRs as substrates, that the absence of thioredoxin s1 in *M. truncatula* leads to an incomplete TBD and that NCRs antimicrobial activity is increased when their cysteines are reduced (Ribeiro et al., 2017).

In *Medicago*, other types of plant-secreted peptides were identified and studied, such as the GRP (Glycine Rich Peptides). GRP are a family of plant secreted peptides known to be involved in various processes such as in biotic and abiotic interactions (Sachetto-Martins et al., 2000). Though plant GRP generally have 80% of their sequence composed of glycines, GRP

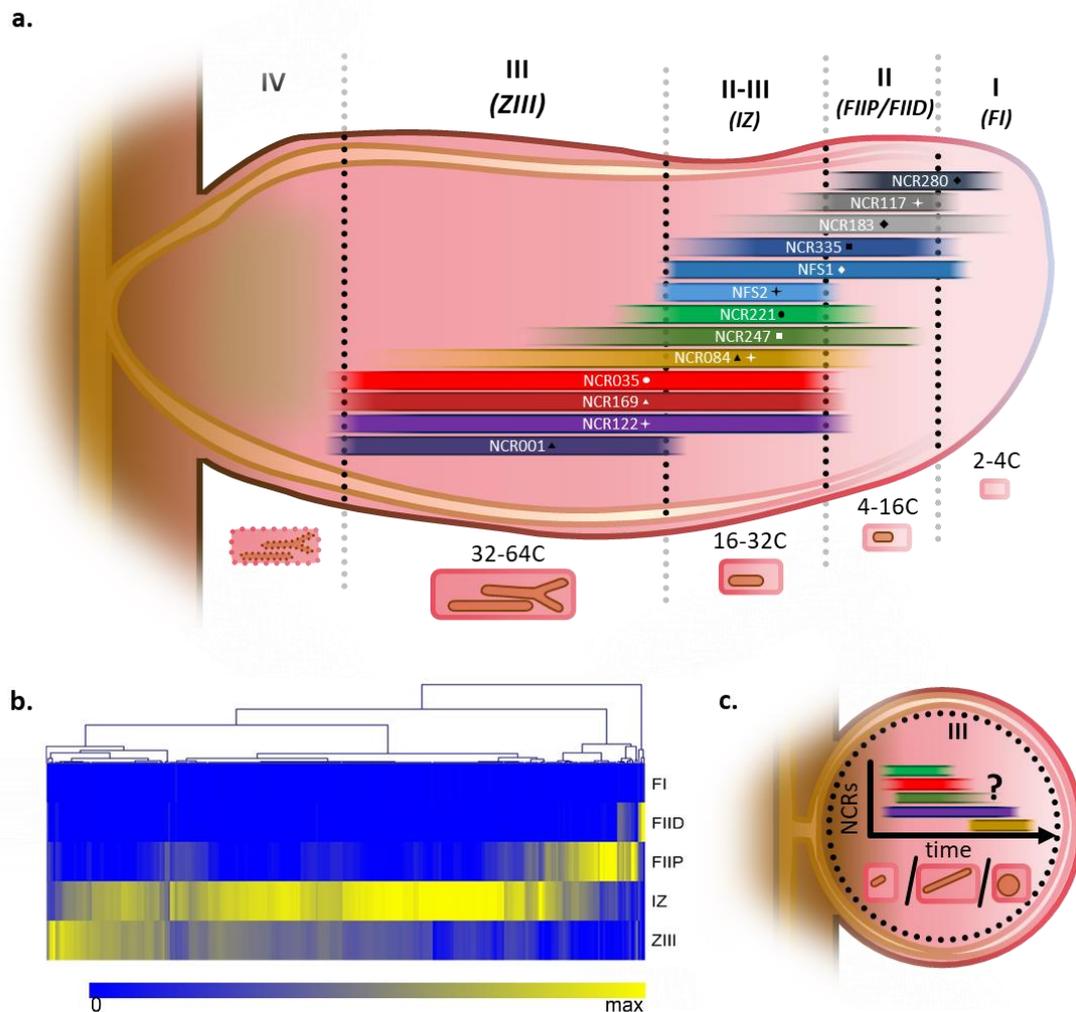


Figure 23 | Expression pattern of NCR genes. a. | Localization of selected NCR peptides in the different zones of *Medicago truncatula* nodule. The zones, in which indeterminate nodules of *Medicago* species are segmented, are written in two equivalent nomenclatures. In one of them, zone II can be separated into ZIID (distal), on the tip side, and ZIIP (proximal), on the root side. For each zone, the maximum ploidy level of plant cells is mentioned. Next to each name of NCR peptides, symbols indicate the study from which the spatial expression was determined. Black lozenge: transcriptomics, Guefrachi et al., 2014; white cross: transcriptomics, Nagyhamly et al., 2017; black square: transcriptomics Ribeiro et al., 2017 and Tiricz et al., 2013; white lozenge: promoter activity, Wang et al., 2017; black cross: promoter activity, Yang et al., 2017; black circle: promoter activity, Kim et al., 2015; white square: promoter activity, Farkas et al., 2014; black triangle: *in situ* hybridization, Mergaert et al., 2003; white circle: translational fusion, Van de Velde et al., 2010; white triangle: translational fusion and promoter activity, Horvath et al., 2015. **b.** | Expression profiles of NCR genes in the different zones. Extracted from Guefrachi et al., 2014, supplemental data. **c.** | In nodules of Dalbergioid plants inducing TBD, few is known on NCR gene expression.

Aeschynomene NCR motif 1	SP-X _n -	C-X _n -C-X ₃ -C-X _n -C-X ₄ -C-X ₁ -C-X _n
Medicago NCR motif a	SP-X _n -	C-X ₅ -C-X _n -C-X _n -C-X ₄ -C-X ₁ -C-X _n
Aeschynomene NCR motif 2	SP-X _n -	C-X _n -C-X _n -C-X ₃ -C-X _n -C-X ₄ -C-X ₁ -C-X ₃ -C-X _n
Defensin-like motif	SP-X _n -	C-X _n -C-X _n -C-X ₃ -C-X _n -C-X _n -C-X ₁ -C-X ₃ -C-X _n

Figure 24| Comparison of *Aeschynomene* and *Medicago* NCR motifs with defensin motifs.
Extracted from Czernic et al., 2015.

which are specifically expressed in nodules have fewer glycines (20-30%) and are shorter. Moreover, GRPs from IRLC show sequence variations which distinguishes them from other nodule GRPs (Alunni et al., 2007; Kevei et al., 2002). Interestingly, 39% of *M. truncatula* GRP transcripts are present in the infection zone whereas only 19% of NCR transcript are represented in this zone in the same plant (Kereszt et al., 2018). It is possible GRP are important for the early steps of the TBD. However, current knowledge on nodule specific GRPs is limited and no functional studies demonstrated the roles of GRPs in nodules.

NCR-like peptides in Dalbergioids

Another legume clade in which TBD occurs is the *Dalbergioid*. A study of nodule transcripts from *Aeschynomene* species led to the identification of NCR-like genes, a new class of the symbiosis controlling peptides (Czernic et al., 2015). Depending on the species, the number of NCR found varied from 37 to 80. Their sequence diverges from the NCRs of *M. truncatula*, even in the organization of the cysteine residues. Indeed, two motifs are present in *Aeschynomene* NCR peptides (**Fig. 24**). One resembles the NCR motif of *Medicago* with six cysteines, while the other has eight of them, therefore being closer to the plant defensin patterns (Czernic et al., 2015). However, their overall sequence is quite divergent, which suggests a convergent evolution process by an independent acquisition of NCRs by IRLC and *Dalbergioids* possibly form the innate immune repertoire. Like *Medicago* NCR genes, the genes encoding the NCR-like peptides were found to be expressed in nodule cells and to colocalize with leghemoglobin in *A. afraspera* or *A. indica*. Proteomics showed that the peptides accumulated in bacteroids. As the targeting of NCR-like peptides to symbiosomes involves the secretion machinery, a *DNF1* homolog of *A. evenia* ssp. *serrulata*, *AeDNF1*, was knocked-down using RNA interference resulting in a strong TBD defect as bacteroids could not reach a spherical morphology (Czernic et al., 2015). As the colonized tissues of indeterminate nodules such as in Dalbergioid clade

have a homogeneous colonized tissue where bacteroids are all in the same state, it is likely that NCR peptides expression is only regulated temporally and not spatially (**Fig. 23c**).

Strikingly, *Bradyrhizobium* sp. ORS285 treated with synthetic versions of some of the most abundant *Aeschynomene* spp. NCR-like peptides showed neither growth defect nor TBD features. Though it seems bradyrhizobia could have been resistant to cationic NCRs due to their membrane properties, these previous results could be explained by the nature of NCR-like peptides (Czernic et al., 2015). Most of detected NCR-like peptides in *Aeschynomene* species are neutral or anionic whereas only cationic NCR peptides were reported so far to have an *in vitro* activity. Some undetected classes of NCR-like peptides might still need to be identified in *Aeschynomene* species to fully understand the TBD in Dalbergioids. Alternatively, the membrane composition of bradyrhizobia could be modified *in planta* and make them sensitive to NCR-like peptides in this specific context.

iii. How TBD impacts symbiosis – The notion of Symbiotic efficiency

Through the action of NCRs, bacteroids are pressed into a metabolically active state but which is nevertheless a dead end. They are fully dedicated to nitrogen fixation but have lost their division ability and capacity to live independently. On the other hand, in many legumes, TBD is not taking place but the bacteroids also function as nitrogen-fixing factories. This raises the important question about the biological meaning of the TBD. By this process, it is believed that the symbiotic efficiency is increased, and some studies corroborate this hypothesis. *Rhizobium leguminosarum* A34 is able to elicit nodules in pea and bean and will differentiate either in TBD-type or non-TBD-type bacteroids, respectively. Pea plants had a greater dry weight per gram of nodule as well as increased nitrogen fixation compared to bean (Oono & Denison, 2010). In the same study, a similar scenario takes place with peanut and cowpea, respectively inducing TBD and non-TBD bacteroids in nodules elicited by *Bradyrhizobium* 32H1 (Oono & Denison, 2010). Thus, comparing symbiosis of the same rhizobium strain in legumes forming swollen or non-swollen bacteroids suggests that the interaction is more efficient when TBD occurs. A more recent study corroborates these findings by showing that there is a positive correlation between the degree of TBD (bacteroid size and ploidy levels) and symbiosis efficiency (Kazmierczak et al., 2017). Indeed, *M. truncatula* can be inoculated by different compatible bacterial strains, which does not display equivalent symbiotic efficiencies. The

stronger the symbionts elongate and accumulate DNA, the more nitrogen fixation and plant biomass increase (Kazmierczak et al., 2017). Lastly, a study comparing S-type and E-type bacteroids in the Dalbergioids found that the S-type bacteroids are more polyploidy and more efficient (Lamouche et al., 2019b). Symbiotic efficiency is often assessed with a group of parameters. Some of them estimate the gain of biomass of the plant. In this case, the shoot, root and nodule masses can be measured and used in various calculations to quantify the symbiotic efficiency. Shoot-root ratio is one of them representing the nutritional status of the plant. Symbiotic efficiency can also be expressed as the mass gain per mg of nodules, where mass gain is calculated as the weight difference between inoculated plants and non-inoculated plants. This last parameter allows inter-species comparisons (Lamouche et al., 2019b). Besides these weight measurements, carbon-nitrogen content or nitrogenase activity can be assessed. All these measurements estimate the return on investment for host plants. However, seed yield is rarely measured in laboratory conditions though being the parameter of choice for agricultural engineers or farmers.

iv. Bacterial response to differentiation – which functions to survive the NCR burst

Envelope functions

Several bacterial functions are known to be involved in resistance to AMPs (**Fig. 25**). It is possible they are also involved in enduring the presence of NCR peptides. Indeed, the NCR burst in nodules can be stressful for bacteroids and they need to resist this challenge to survive. The first protection bacteria could build against NCR is an adapted envelope as cationic NCRs directly target it. Here we are going to follow the path of NCRs peptides from the peri-bacteroid space to the bacteroid cytosol. The most external protection a bacterial envelope can harbor is an EPS coating. With their mesh structure, EPS could slow down the diffusion of NCR peptides, helping bacteria resisting NCRs (**Fig. 25**) (Arnold et al., 2018; Montiel et al., 2017). The second layer the AMPs reach is the LPS which is known to be involved in avoiding defense reaction in host-bacteria interactions. However, it is also known that they are involved in resistance to some AMPs as well. Thanks to their O-antigen, LPS are known to protect bacteria from the action of cationic AMPs by masking the negative charge of the bacterial

membrane. This protection reduces the affinity of cationic AMPs for the bacterial envelope (**Fig. 25**) (Campbell et al., 2002). Various mutants that affect the structure of LPS show symbiotic deficiency, though being able to elicit and infect nodules. None of them was directly linked to NCR activity yet though. For example, the absence of the *lpsB* gene in *S. meliloti*, leads to an early senescence. Indeed, this gene is involved in the synthesis of the LPS-core and possesses

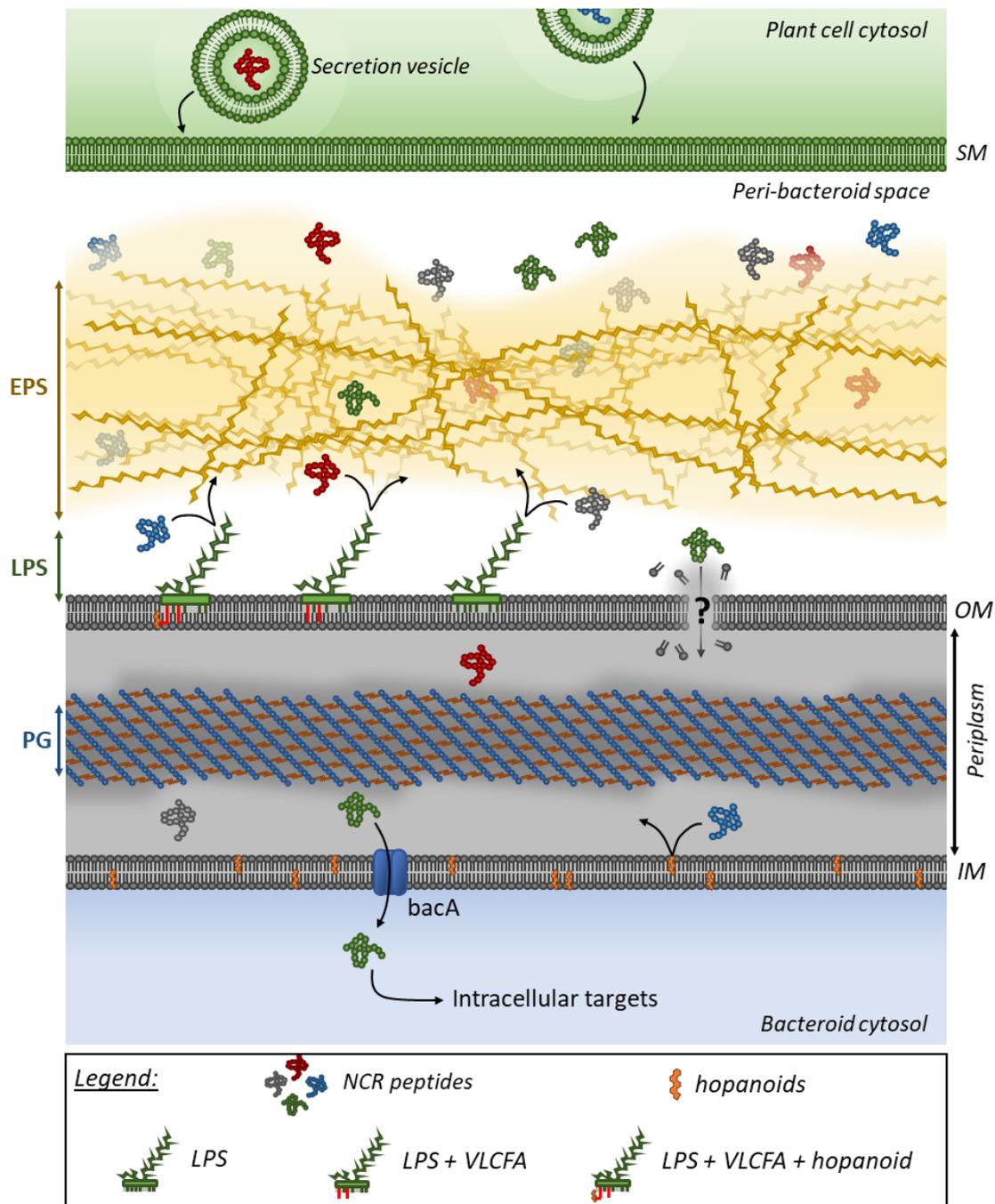


Figure 25 | Envelope functions of bacteroids potentially involved in NCRs. This figure schematizes the barriers bacteria can set up in order to resist the NCR burst. Scales are not respected. EPS: exopolysaccharides; IM: inner membrane; LPS: lipo-polysaccharides; OM: outer membrane; PG: peptidoglycan; SM: symbiosome membrane.

altered LPS (Campbell et al., 2002). *acpXL* and *lpxXL* mutants are affected in the synthesis of VLCFA and in the attachment to the LPS lipid A respectively. VLCFA are synthesized in the cytosol, linked to the AcpXL (Acyl Carrier Protein) protein. Afterwards, they are transferred to the lipid A thanks to the membrane protein LpxXL (Long chain acyltransferase) (Haag et al., 2011; Sharypova et al., 2003). Without *acpXL*, nodules prematurely senesce while without *lpxXL* bacteroids have abnormal morphology (Haag et al., 2009). Mutation of the orthologous genes in *R. leguminosarum* bv. *viciae* 3841 led to delayed infection, early senescence, and sensitivity to AMPs (Bourassa et al., 2017). Similar results were also observed in the absence of *lpxXL* homologs in *Bradyrhizobium* sp. ORS278 (Busset et al., 2017).

Sterol-like hopanoids have also been shown to be important for membrane stability. Those lipids can either be linked to the VLCFA attached to the LPS of some *Bradyrhizobium* strains (for some C35 hopanoids) or be free in the inner membrane (all other hopanoids) (**Fig. 25**) (Busset et al., 2017; Kulkarni et al., 2015; Silipo et al., 2014). Absence of hopanoids leads to an increased sensitivity to membrane stresses, loss of membrane integrity in nodules and ultimately a plant growth defect (Silipo et al., 2014). Particularly, the function of two types of hopanoids was further investigated in symbiosis in *B. diazoefficiens* USDA110. On one side, the C35 hopanoids, which rigidify membranes, are synthesized by HpnH. On the other side, 2Me-hopanoids are produced by HpnP particularly in stress conditions (Kulkarni et al., 2015). Both forms are important for growth in microaerobic conditions. In addition, a *hpnH* mutant displays increased sensitivity to membrane stress, including to the NCR335 peptide from *M. truncatula*. Interestingly, this mutant displayed an inefficient interaction with *A. afraspera* but no defects in its interaction with soybean (Kulkarni et al., 2015). Bacteroid membrane breakdown and induction of plant defense reactions were only observed in *A. afraspera* nodules. It is possible that this difference is due to the presence NCR-like peptides in *A. afraspera*. In this case, hopanoids could help bacteria to resist the NCR stress thanks to their membrane rigidifying properties.

NCR resistance also requires a functional peptidoglycan. This complex polymer composed of saccharides and small peptides is essential in the structuration of the envelope and in maintaining the shape of the bacteria (Typas et al., 2012; van Teeseling et al., 2017). Various genes involved in processing peptidoglycan were shown to be essential to survive a

NCR247 treatment in *S. meliloti* 1021 (**Fig. 25**) (Arnold et al., 2017). Moreover, a functional study of a peptidoglycan remodeling enzymes showed strong impairment of the symbiotic process in *Bradyrhizobium* species (Gully et al., 2016). A mutant in the *DD-Cpase1* gene of *Bradyrhizobium* sp. ORS278 was only able to elicit badly colonized or empty nodules on *A. indica*. Plant defense responses were observed in these nodules. Consequently, the plants grew smaller than the ones infected by the WT strain. The few bacteroids that were present in infected cells were gigantic with an aberrant cell shape and displayed an increased DNA content, as compared to WT bacteroids. These phenotypes were reproduced in *A. afraspera* nodules (Gully et al., 2016). However, the mutation of the orthologous gene in *B. diazoefficiens* USDA110 did not alter bacteroid shape in *G. max*. Therefore, it was suggested that the symbiotic function of this peptidoglycan modifying enzyme is related to the presence of NCR-like peptides in nodules.

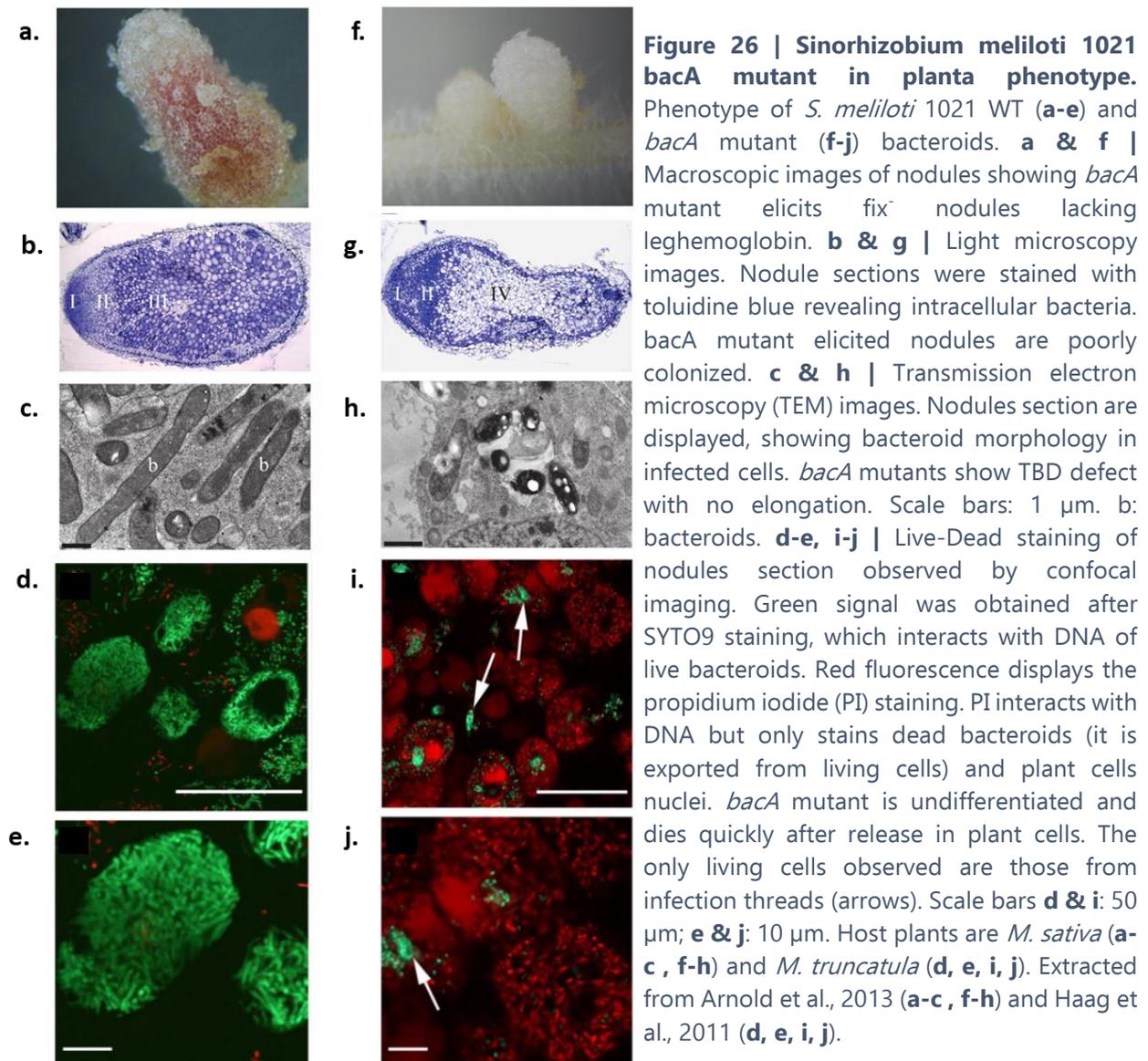
It should be noted that although these envelope functions affect the resistance of bacteroids to the antimicrobial NCRs, other types of stresses, like acidic pH, low oxygen or reactive oxygen species, are faced by bacteroids in nodules (Gibson et al., 2008; LeVier et al., 2000; Perez-Galdona & Kahn, 1994). It is possible that the bacterial envelope functions also contribute to handle these additional stress conditions in the nodule environment.

Transporting NCRs to protect the envelope – The role of BacA

NCR peptides have both antimicrobial and TBD inducing activities. Both activities can be obtained by treating *S. meliloti* cultures with purified NCR peptides (Van de Velde et al., 2010). Bacteroids need to keep NCRs away from their envelope to protect it. In the previous section bacterial functions were discussed, used to strengthen or to conceal the envelope. In addition, transporters also seem to be involved in this function (**Fig. 25**). The BacA transporter was identified thirty years ago, before the identification of NCR peptides because a mutant in the gene is only able to elicit fix⁻ nodules (Glazebrook et al., 1993). The *bacA* gene encodes a membrane protein resembling an ABC transporter. Import transporters from this family are usually made of three sub-units. The most external sub-unit is called the Periplasmic-binding Protein (PBP) present in the periplasm. It binds the transporter ligand and transfers it to the permease domain. This transmembrane sub-unit is usually composed of two proteins localized on the inner membrane. To carry out its activity, the permease requires energy provided by

ATP hydrolysis thanks to the ATPase domain present in the cytosol (Davidson et al., 2008). BacA has a different structure as it has no ATPase domain, and no PBP could be associated to this transporter. Its permease domain is probably a homodimer (Corbalan et al., 2013; LeVier & Walker, 2001; Runti et al., 2013). Considering the absence of an ATPase domain to energize the transporter, BacA probably requires an electrochemical membrane gradient to be functional as shown for its homolog SbmA in *E. coli* (Runti et al., 2013).

During symbiosis, *bacA* is expressed in the infection (II), differentiation (II-III), and fixing (III) zones (Glazebrook et al., 1993; Roux et al., 2014). In *M. truncatula* or *M. sativa*, the *bacA* mutant induces nodules and infects them through infection threads, but bacteria die as soon as they are released into the nodule cells (**Fig. 26**). Although the transcriptome of $\Delta bacA$ bacteroids is highly similar to WT bacteroids at the early states, at 11dpi they become quite



different corresponding to the absence of differentiation of the symbiont in plant cells (Capela et al., 2006). Strikingly, the BacA transporter is required only in NCR-producing plants. A study of the *L. japonicus* and *M. loti* interaction showed the *bacA* mutant of *M. loti* induces functional nodules. Though being less efficient compared to nodules elicited by WT strain (reduction of 20% to 50% of nitrogen fixation), plant growth is not affected (Maruya & Saeki, 2010). This phenotype was also reported in bean (Karunakaran et al., 2010) and other non-IRLC and non-Dalbergioid legumes (Ardissone et al., 2011).

BacA has homologs in numerous bacteria, including human and animal pathogens. It has been proposed that these homologs are important to protect bacteria from host defense reaction (LeVier et al., 2000) and their importance for the persistence of intracellular infection was later confirmed (Domenech et al., 2009). BacA's homolog SbmA was shown to import a glycine-rich AMP, microcin B17, which interferes with the DNA replication machinery. In absence of SbmA, *E. coli* becomes more resistant to microcin B17. Several studies show that it is not the only AMP imported by SbmA (Mattiuzzo et al., 2007). From all these data was deduced that SbmA can transport cationic peptides. Thanks to cross-complementation, the functional homology of SbmA and *S. meliloti* BacA has been demonstrated (Ichige & Walker, 1997). Indeed, an over-expression of *bacA* in *E. coli* $\Delta sbmA$, restores the sensitivity of the strain to microcin J25, B17 and other peptides which require an internalization for their toxic activity. Reciprocally, expressing *sbmA* in *S. meliloti* $\Delta bacA$ strains suppresses every phenotype resulting from the mutation. Even the symbiotic defects are suppressed (Arnold et al., 2014; Ichige & Walker, 1997; Marlow et al., 2009). To corroborate these results, it was shown that the presence of BacA increases sensitivity to the peptide Bac7. This AMP, which name has no relation with the transporter BacA, interacts with ribosomes and inhibits translation (Mardirossian et al., 2014; Seefeldt et al., 2016). Removal of BacA prevents internalization of the peptide and therefore its antimicrobial activity (Marlow et al., 2009). Taken together, these data strongly suggest BacA is an AMP importer.

However, the *bacA* mutant phenotypes might not only be due to the AMP transport activity of BacA. Indeed, it has been shown that the absence of BacA leads to increased sensitivity to various membrane stresses (acidity, detergents, ethanol, and aminoglycoside antibiotics) (Ichige & Walker, 1997; LeVier & Walker, 2001). This phenotype suggests the

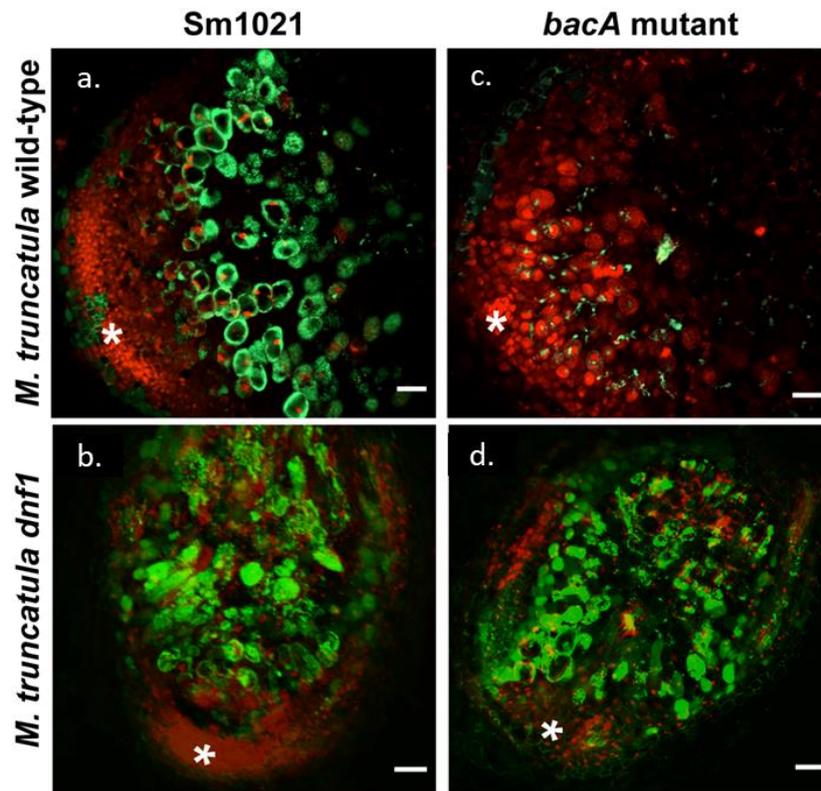


Figure 27 | *Sinorhizobium meliloti* 1021 *bacA* mutant phenotype within *M. truncatula dnf1* nodules. Confocal microscopy after Live-Dead staining reporting the viability of symbiosome bacteria in the *M. truncatula dnf1* mutant nodules. Stains used are SYTO9 (green) and PI (red). The host plant is either WT *M. truncatula* (a & c) or *dnf1* mutant (b & d). Symbiont are either the WT *S. meliloti* 1021 strain (a & b) or *bacA* mutant (c & d). Whereas *S. meliloti* WT survives in *dnf1* nodules, with no TBD though, the *bacA* mutant, which usually dies quickly in plant cells, is able to survive in *dnf1* nodules, suggesting that NCR peptides kill the *bacA* mutant in WT nodules. The nodule meristem is indicated by an asterisk. The strong red PI staining in the meristem is resulting from the high density of plant nuclei. Scale bars: 50 μ m. Extracted from Haag et al., 2011.

composition of the membrane is different between the *bacA* mutant and its parent strain. It was found that LPS lipid A is different as some of the VLCFA were absent from the lipid A moiety in the mutant (Ferguson et al., 2002, 2004). To assess whether $\Delta bacA$ phenotype is due to this alteration of LPS structure, the role of VLCFA synthesis genes in symbiosis was analyzed. However, the *acpXL* and *lpxXL* mutants did not show as strong symbiotic phenotype as the *bacA* mutant. Therefore, even if the *bacA* phenotype is partly due to the alteration of lipid A, its AMP-transporting activity also plays part in its symbiotic phenotype. Moreover, in the study of Bac7 effects on *bacA* mutants, it was found that the *acpXL* mutant, which has no VLCFA, has increased sensitivity to Bac7. As *bacA* has the opposite phenotype, this phenotype is independent of LPS modifications.

Based on these data, it was proposed that BacA is involved in NCR peptide recognition and import (Mergaert et al., 2006). Subsequently, it was indeed shown that this protein is important for the resistance of bacteria to NCR035 and NCR247 *in vitro* treatments (Arnold et al., 2017; Haag et al., 2011). These experiences show that *bacA* is required *in vitro* to resist lytic NCR peptides. However, the *bacA* gene seems not to be required to induce the TBD phenocopy *in vitro*. Treatment with 4 μM of NCR247 is enough to induce increase of DNA content in WT and *bacA* mutant equally. A possible explanation lies in the membrane-penetrating activity of the cationic peptide, which would allow the NCR peptide to enter bacteria anyway. However, the results that strongly link *bacA* to NCRs *in planta* are the analysis of the *bacA* mutant in interaction with *dnf1* plants (**Fig. 27**). The mutant bacteroids survive in *dnf1* nodules contrary to wild type nodules where the mutant dies immediately after release in the nodule cells (Haag et al., 2011;).

bacA has also a homolog in *Bradyrhizobium* sp. ORS278 and ORS285 called *bclA* for “*bacA*-like *A*”. This gene differs slightly from *bacA* as it possesses an ATPase domain. Similarly, it is essential for the differentiation of bacteroids in the NCR-producing *A. indica* and *A. afraspera* nodules (Guefrachi et al., 2015). The analysis of the DNA content in bacteroids by flow cytometry revealed that the *bclA* mutant cannot differentiate in nodules from these host plants. Nitrogen fixation is strongly impaired as well. However, phenotypes of the *bclA* mutant slightly differs from one host to the other. Bacteroids in *A. afraspera* are all alive and low levels of nitrogen fixation could be detected. In *A. indica*, there is a mixed population of dead and live bacteroids at 14 dpi. Bacteroid death happens progressively in the nodules (Guefrachi et al., 2015). Similarly, to *bacA* mutants in *Rhizobium* or *Mesorhizobium* strains in plants that do not produce NCR peptides, *bclA* seems not to be required for symbiosis between *B. diazoefficiens* USDA110 and soybean (Barrière et al., 2017). Roles or functioning of *bacA* and *bclA* *in planta* slightly differ as ORS285 *bclA* can complement the loss of *S. meliloti* *bacA* but *bacA* cannot complement $\Delta bclA$ (Barrière et al., 2017; Guefrachi et al., 2015). This feature is likely linked to the structure of both orthologs, more precisely the presence-absence of the ATPase domain, and thus different means of energization of the transporter. As BacA should require a third-party factor to provide energy or an electrochemical gradient, it is possible that such factor is absent in *Bradyrhizobium* spp. as BclA has the ATPase domain for energy.

Some bacteria totally resist the NCR burst

Strikingly, TBD in IRLC has been shown to be avoided by some symbionts. These bacteria developed strategies to escape the NCR selection. As an example, *S. fredii* HH103 can interact with *Glycyrrhiza uralensis*, an IRLC legume. Though this plant produces NCR peptides (Montiel et al., 2017), HH103 bacteroids display neither cell elongation nor polyploidization (Crespo-Rivas et al., 2016). In this strain, *bacA* is not required for a functional symbiosis. It is possible that the O-antigen from LPS, which has an altered structure compared to *S. meliloti* 1021, helps protecting the strain from NCR peptide activity. Additionally, it has been shown that EPS and KPS structure are quite different between *S. meliloti* 1021 and *S. fredii* HH103 (Frayse et al., 2005; Fraysse et al., 2003; Margaret et al., 2011; Rodríguez-Navarro et al., 2014). Some other *S. meliloti* strains, which possess the *hrrP* (host range restriction peptidase) gene on an accessory plasmid, also display NCR resisting features. Depending on the host plant with which they interact, they can display either fix^- or fix^+ phenotype (Price et al., 2015). This peptidase can cleave NCR peptides *in vitro*, and strains expressing HrrP induce an early nodule senescence. The corresponding bacteroids may undergo a kind of undifferentiation (or most probably a milder differentiation) and can recover growth when extracted from nodules (Price et al., 2015).

E. Objectives of the thesis project

My thesis work was articulated around the TBD triggered by NCR peptides during the legume–rhizobium symbiosis. Nearly fifteen years ago, NCR peptides and their TBD inducing activity were discovered in *Medicago*. Since then, extensive research was conducted both on the plant side, to appraise the diversity of IRLC NCR peptides in the plants or their secretion mechanisms, and on the bacteria to address their effects on the bacteroid life. Localization and activity of several cationic NCR peptides was assessed. Five years ago, a new class of NCR peptides was discovered in Dalbergioids, which functions have been shown to overlap with IRLC NCR peptides. However, our current knowledge on their activity and the bacterial response is limited. During my PhD, I studied different legume – rhizobium symbiotic models to better understand the NCR regulation of the TBD as well as the bacterial response to this stress.

Dalbergioids–rhizobium symbioses display quite remarkable features such as Nod factor-independent initiation processes, stem nodulation or spherical bacteroids. Next to IRLC,

Dalbergioid is the only clade that acquired NCR peptides for symbiont controlling purposes. Interestingly, *B. diazoefficiens* USDA110, which usually interacts with soybean (*G. max*), can perform a functional symbiosis with the Dalbergioid species *Aeschynomene afraspera*. However, in this symbiosis, bacteroids are not elongated, contrary to *Bradyrhizobium* sp. ORS285 bacteroids in *A. afraspera*. I studied these two interactions with two purposes. On one side, we wanted to fully assess USDA110 bacteroid behavior in nodule cells in order to better understand how this symbiosis can be functional. On the other side, we conducted multi-omics analyses, comparing the two models and the natural symbiosis of USDA110 to understand how this bacterium adapts to *A. afraspera* and to *G. max*.

Although the roles of *Medicago* NCR peptides have been intensively studied, at least for the cationic ones, the corresponding bacterial responses are poorly known today. With the whole battery of NCR peptides *Medicago* possesses and given the complexity and large array of bacterial features that are altered in bacteroids, there are surely numerous bacterial functions that are directly or indirectly affected by NCR peptides. Until now, it is known that BacA is an ABC transporter involved in NCR import, but its phenotype alone does not explain the whole TBD process. Indeed, the way NCR peptides trigger TBD is still blurry. To better understand the functions required by bacteroids to survive the activity of NCRs, we assessed the importance of different bacterial genes for their involvement in the NCR response. Functional analysis of a set of *S. meliloti* mutants was conducted addressing their role in symbiosis and whether they are linked with NCR activities.

THESIS PROJECT STUDIES

A. *Bradyrhizobium diazoefficiens* USDA110 nodulation of *Aeschynomene afraspera* is associated with atypical terminal bacteroid differentiation and suboptimal symbiotic efficiency.

Authors : Quentin Nicoud^{1*}, Florian Lamouche^{1*}, Anaïs Chaumeret¹, Thierry Balliau², Romain Le Bars¹, Mickaël Bourge¹, Fabienne Pierre¹, Florence Guérard³, Erika Sallet⁴, Solenn Tuffigo¹, Olivier Pierre¹, Yves Dessaux¹, Françoise Gilard³, Bertrand Gakière³, Istvan Nagy^{5,6}, Attila Kereszt^{5,6}, Michel Zivy², Peter Mergaert¹, Benjamin Gourion⁴, Benoit Alunni¹

¹ Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, 91198, France

² PAPPISO, GQE-Le Moulon, INRAE, CNRS, AgroParisTech, Paris-Saclay University, 91190 Gif-sur-Yvette, France

³ SPOmics platform, Institute of Plant Sciences Paris-Saclay (IPS2), CNRS, INRAE, Universities Paris-Saclay, Evry and de Paris, Batiment 630, 91405 Orsay, France

⁴ LIPM, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France

⁵ Institute of Biochemistry, Hungarian Academy of Sciences, Biological Research Centre, 6726 Szeged, Hungary

⁶ Seqomics Biotechnology Ltd., 6782 Mórahalom, Hungary

(* : equal contribution)

i. Foreword

In this project, we studied the ill-adapted interaction of *B. diazoefficiens* USDA110 with *A. afraspera*. Indeed, USDA110 is the natural symbiont of soybean and this plant does not produce NCR peptides. Therefore, USDA110 is neither adapted to this stress nor to undergo the TBD process orchestrated by NCR peptides. Interestingly, USDA110 is still able to infect NCR-

producing *A. afraspera* plants (Renier et al., 2011). These are striking features, as usually, bacteria that are not able to handle the NCR stress, such as *bclA/bacA* mutants, quickly die in the nodules, which in turn undergo early senescence. *B. diazoefficiens* USDA110 is known to have an envelope that is highly resistant to membrane stresses. Notably the production of hopanoids rigidifies its envelope, which could explain its survival in nodules. However, USDA110 is not able to elongate *in planta*. Moreover, no increase in DNA content is observed in this strain, suggesting that USDA110 can resist the TBD.

Two main goals drove this study. On one side, we aimed to study the molecular functions that underlie bacteroid differentiation, which is linked to NCR-like peptides. In that scope, USDA110 bacteroids from *A. afraspera* were used as a tool to mimic the behavior of NCR-resistant bacteroids in this plant. By comparing this interaction to the USDA110 and soybean symbiosis, we assessed the functions necessary to the symbiont life in an unusual host plant and in presence of NCR-like peptides. Comparing USDA110 bacteroids to *Bradyrhizobium* sp. ORS285 bacteroids from *A. afraspera* (previously studied by Lamouche et al., 2019) allowed us to compare the functions required to a host-adapted symbiont, which is used to respond to the NCR-like burst with those of the un-adapted symbiont. To do so, we conducted a multi-omics approach combining transcriptomics, proteomics, and metabolomics. Moreover, this study helped us understand how an ill-adapted bacterium can survive an unnatural host. On the other side we tried to better dissect the phenotype of USDA110 bacteroids in *A. afraspera* by investigating various TBD features. This analysis comprises an assessment of symbiotic efficiency of these bacteroids in the two host plants.

This work began well before my PhD started, with the generation and analyses of bacteroid transcriptomes, proteomes, and metabolomes. I carried out the comparison of ortholog genes behavior in transcriptomic and proteomic data as well as an analysis of differentially expressed genes (DEG) and differentially accumulated proteins (DAP). I designed and conducted the analyses of TBD features such as membrane permeability assays and survival rate determination. Together with Romain Le Bars, we conducted the morphometric analysis of bacteroid shape. The paper was submitted to the mSystems journal in late November 2020 and can be found on the Biorxiv platform at the following link: <https://biorxiv.org/cgi/content/short/2020.11.24.397182v1>.

ii. Abstract

Legume plants can form root organs called nodules where they house intracellular symbiotic rhizobium bacteria. Within nodule cells, rhizobia differentiate into bacteroids, which fix nitrogen for the benefit of the plant. Depending on the combination of host plants and rhizobial strains, the output of rhizobium-legume interactions is varying from non-fixing associations to symbioses that are highly beneficial for the plant. *Bradyrhizobium diazoefficiens* USDA110 was isolated as a soybean symbiont but it can also establish a functional symbiotic interaction with *Aeschynomene afraspera*. In contrast to soybean, *A. afraspera* triggers terminal bacteroid differentiation, a process involving bacterial cell elongation, polyploidy and membrane permeability leading to loss of bacterial viability while plants increase their symbiotic benefit. A combination of plant metabolomics, bacterial proteomics and transcriptomics along with cytological analyses was used to study the physiology of USDA110 bacteroids in these two host plants. We show that USDA110 establishes a poorly efficient symbiosis with *A. afraspera*, despite the full activation of the bacterial symbiotic program. We found molecular signatures of high stress levels in *A. afraspera* bacteroids whereas those of terminal bacteroid differentiation were only partially activated. Finally, we show that in *A. afraspera*, USDA110 bacteroids undergo an atypical terminal differentiation hallmarked by the disconnection of the canonical features of this process. This study pinpoints how a rhizobium strain can adapt its physiology to a new host and cope with terminal differentiation when it did not co-evolve with such a host.

iii. Importance

Legume-rhizobium symbiosis is a major ecological process in the nitrogen cycle, responsible for the main input of fixed nitrogen in the biosphere. The efficiency of this symbiosis relies on the coevolution of the partners. Some legume plants, but not all, optimize their return-on-investment in the symbiosis by imposing on their microsymbionts a terminal differentiation program that increases their symbiotic efficiency but imposes a high level of stress and drastically reduce their viability. We combined multi-omics with physiological analyses to show that the non-natural symbiotic couple formed by *Bradyrhizobium diazoefficiens* USDA110 and

Aeschynomene afraspera is functional but displays a low symbiotic efficiency associated to a disconnection of terminal bacteroid differentiation features.

iv. Introduction

Nitrogen availability is a major limitation for plant development in many environments, including agricultural settings. To overcome this problem and thrive on substrates presenting a low nitrogen content, crops are heavily fertilized, causing important environmental damage and financial drawbacks (1,2). Plants of the legume family acquired the capacity to form symbiotic associations with soil bacteria, the rhizobia, which fix atmospheric nitrogen for the plants' benefit. These symbiotic associations lead to the development of rhizobia-housing root organs called nodules. In these nodules, the rhizobia adopt an intracellular lifestyle and differentiate into bacteroids that convert atmospheric dinitrogen into ammonia and transfer it to the plant. Critical recognition steps occur all along the symbiotic process and define the compatibility of the plant and bacterial partners (3). While the mechanisms involved at the early stages of the symbiosis are well described, those of the later stages are much less clear and might affect not only the ability to interact but also the efficiency of the symbiosis (*ie.* the plant benefit).

Nodule-specific Cysteine-Rich (NCR) antimicrobial peptides produced by legumes of the Dalbergioids and Inverted Repeat Lacking Clade (IRLC) were proposed to play a crucial role in the control of host-symbiont specificity at the intracellular stage of the symbiosis (4). NCR peptides are targeted to the bacteroids where they govern the bacteroid differentiation (5-9). In these legumes, the differentiation process entails such profound changes that they suppress the bacteroids' capacity to resume growth and is therefore referred to as terminal bacteroid differentiation (TBD). TBD contrasts with bacteroid formation in legumes that lack NCR genes (*eg.* soybean), where bacteroids are in a reversible state and can resume growth when released from nodules (10). Specifically, TBD is associated with cell elongation, an increase in bacteroid DNA content through a cell cycle switch toward endoreduplication (6,9,11). Furthermore, an increased permeability of the bacteroid envelope also occurs during TBD, most probably due to the interaction of NCR peptides with bacterial membranes (6,7,10,12). Together, these alterations of bacteroid physiology are associated to a strongly decreased viability of the differentiated bacteria, that fail to recover growth when extracted from nodules (6).

While many rhizobia have a narrow host range, some species can nodulate a large array of plant species. One of them, *Bradyrhizobium diazoefficiens* USDA110, can trigger functional nodules without TBD on soybean (*Glycine max*), cowpea (*Vigna unguiculata*) or siratro (*Macroptilium atropurpureum*) (**Fig. 28a-b**) (13). In addition to these species, USDA110 induces also functional nodules on the TBD-inducing legume *Aeschynomene afraspera* (**Fig. 28a-c**) (14,15). However, in *A. afraspera*, USDA110 shows only very limited features that are usually associated with TBD, suggesting that the bacterium might be resistant to the TBD process (16).

Herein, we further characterized the bacteroid differentiation in the symbiosis between USDA110 and *A. afraspera*. Our observations, supported by whole-nodule metabolome analysis, indicate that USDA110 is poorly matched for nitrogen fixation with *A. afraspera*. To understand better the adaptation of *B. diazoefficiens* physiology to the *G. max* and *A. afraspera* nodule environment, we used a combination of transcriptomics (RNA-seq) and shotgun proteomics (LC-MS/MS) approaches. Finally, we find that USDA110 undergoes a terminal but atypical bacteroid differentiation in *A. afraspera* with a reduced cell viability and an increased membrane permeability, while cell size and ploidy level remain unchanged.

v. Results

B. diazoefficiens USDA110 is poorly matched with *A. afraspera* for nitrogen fixation

Previous reports indicate that *B. diazoefficiens* USDA110, the model symbiont of soybean, is able to establish a functional nitrogen-fixing symbiosis with *A. afraspera*, a phylogenetically distant host belonging to the Dalbergioid clade, which naturally interacts with photosynthetic rhizobia such as *Bradyrhizobium* sp. ORS285 (**Fig. 28a-c**) (14-18). To evaluate the efficiency of this symbiosis, nitrogenase activity of USDA110- and ORS285-infected plants and their nitrogen content were determined. Although nitrogenase activity was detected in both types of nodules, it was significantly lower in USDA110-nodulated plants (**Fig. 28d**). A similar trend is observed for mass gain per nodule mass although this difference is not significant (**Fig. 28e**). Nitrogen and carbon contents seemed also reduced in USDA110-nodulated plants as compared to ORS285-nodulated plants, reaching levels similar to those of non-inoculated

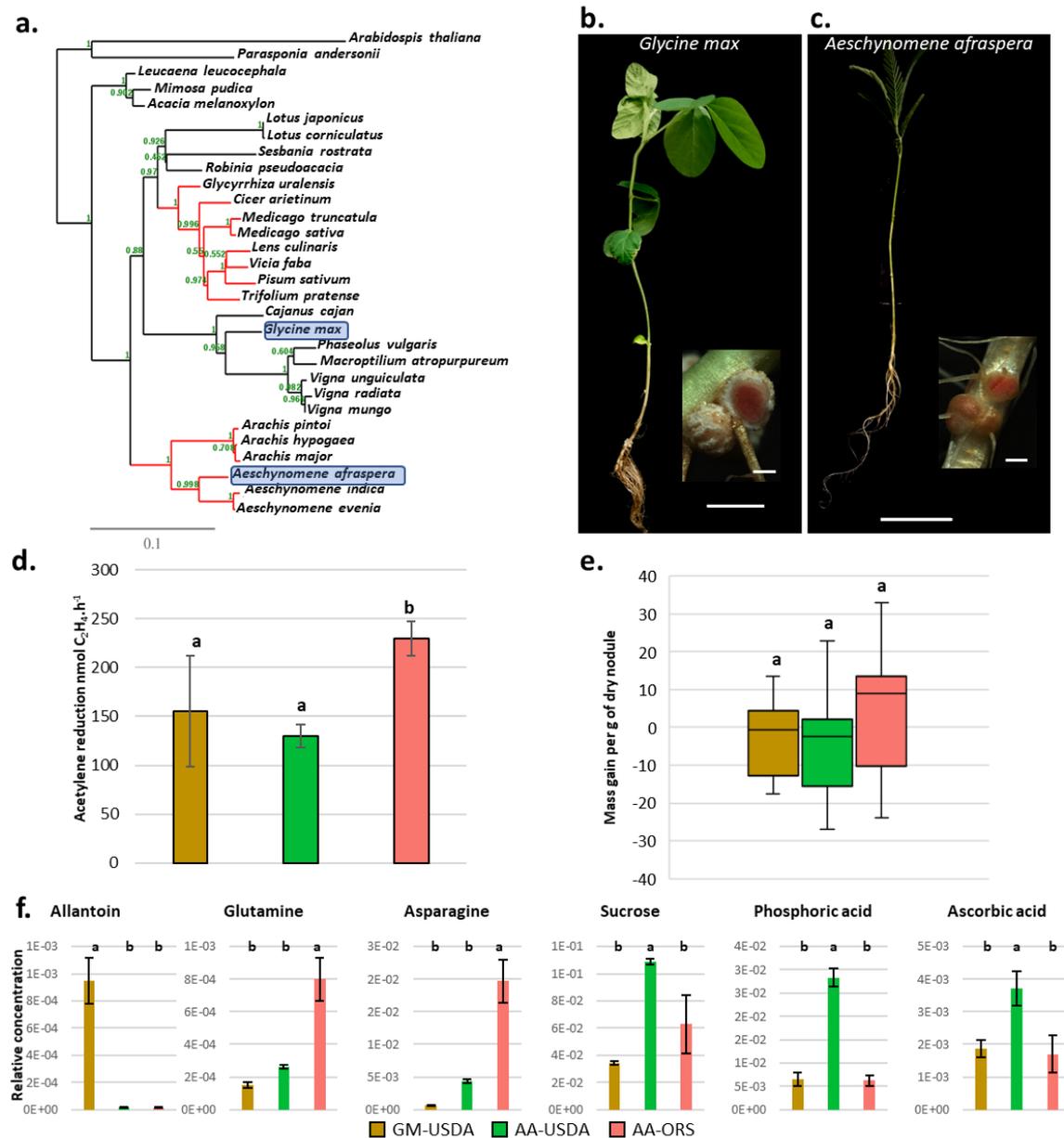


Figure 28 | The non-adapted symbiotic couple formed by *B. diazoefficiens* USDA110 and the NCR-producing plant *A. afraspera* displays suboptimal nitrogen fixation and nodule metabolic dysfunction. **a** | Phylogenetic ML tree of a selection of plant species based on *matK* nucleotide sequences. Red branches indicate clades of legumes plants inducing terminal bacteroid differentiation. Blue boxes indicate the distantly related host plants used in this study. Bootstrap values are mentioned in green on each node of the tree. **b & c** | General aspect of the plants and nodule sections (inlays) displaying the red coloration of leghemoglobin of *G. max* (b) and *A. afraspera* (c) at 14 dpi. Scale bars: 5 cm (plants) and 1 mm (nodules). **d & e** | Nitrogen fixation activity determined by acetylene reduction assay (d) and gain in biomass attributable to the symbiosis (e) of 14 dpi plants. **f** | Whole-nodule metabolome determined by GC/MS or LC/MS at 14 dpi. Histograms show the average value of the relative metabolite concentration of four biological replicates. Letters represent significant differences after ANOVA and post hoc Tukey tests ($p < 0.05$). GM: *G. max* bacteroids, AA: *A. afraspera* bacteroids, USDA: *B. diazoefficiens* USDA110, ORS: *Bradyrhizobium* sp. ORS285.

plants (**Supp. Fig. A1**). Accordingly, ORS285-nodulated *A. afraspera* plants display darker green leaves than those interacting with USDA110.

Moreover, their shoot/root mass ratio, a metrics that reflects the nutritional status of the plant, is reduced in USDA110-nodulated *A. afraspera* plants as compared to ORS285-nodulated plants, indicating that the plant nutritional needs are not fulfilled (**Supp. Fig. A2**) (19). To characterize further this suboptimal symbiosis, we analyzed the whole-nodule metabolome. Soybean nodules infected with USDA110 were used as a reference (**Supp. Fig. A3**). Allantoin, which is known to be the major nitrogen form exported from soybean nodules, is specifically detected in them (**Fig. 28f**) (20). On the contrary, asparagine and glutamine are the principal exported nitrogen compounds in *A. afraspera* nodules and their amount is lower in USDA110-infected nodules as compared to ORS285-infected nodules, indicating a reduced nitrogen fixation by the bacteroids (**Fig. 28f**) (18).

In addition, we find specifically in USDA110-infected *A. afraspera* nodules the accumulation of sucrose, phosphoric acid, and ascorbate, and oppositely, a strong reduction in the trehalose content (**Fig. 28f, Supp. Fig. A3**). Sucrose derived from phloem sap is metabolized in nodules to fuel the bacteroids with carbon substrates, usually dicarboxylates. The accumulation of sucrose in nodules indicates a symbiotic dysfunction. Also, the accumulation of phosphoric acid in nodules suggests that nitrogen fixation is not reaching its optimal rate. Ascorbate has been shown to increase nitrogen fixation activity by modulating the redox status of leghemoglobin (21,22). Thus, its accumulation in nodules with reduced nitrogen fixation capacity could be a stress response to rescue nitrogen fixation in nodules that do not fix nitrogen efficiently. A trehalose biosynthesis gene is upregulated in ORS285 bacteroids in *A. afraspera*, suggesting that TBD is accompanied by the synthesis of this osmo-protectant disaccharide (17). The lower synthesis in USDA110 bacteroids suggests an altered TBD. Together these data indicate a metabolic disorder in the USDA110-infected nodules, in agreement with USDA110 being a suboptimal symbiont of *A. afraspera*.

Overview of the USDA110 bacteroid proteomes and transcriptomes

In order to better understand the poor interaction between USDA110 and *A. afraspera*, the bacteroid physiology was analyzed through transcriptome and proteome analysis. The efficient

> **Figure 29 | Experimental setup and general description of the transcriptomics and proteomics dataset.** **a** | Experimental setup displaying the three biological conditions of this study. **b** | Principal component analysis of the proteomics and transcriptomics datasets. **c** | Venn diagram representing the overlap between differentially expressed genes (DEGs, FDR < 0.01 & |LFC| > 1.58) and differentially accumulated proteins (DAPs, FDR < 0.05) in at least one comparison and among the population of detected proteins. **d** | Pearson correlation coefficient (r) distribution between transcriptomic and proteomic datasets based on DAPs only (red) or DAPs that are also DEGs (green). **e** | Heatmaps and hierarchical clustering of the 815 DAPs and the corresponding transcriptomic expression values. The heatmaps show the standard score (z-score) of assigned spectra and DESeq2 normalized read counts, respectively. The color-coded scale bars for the normalized expression and value of Pearson correlation coefficient of the genes are indicated below the heatmap. YM: Yeast-Mannitol culture, GM: *G. max* bacteroids, AA: *A. afraspera* bacteroids.

soybean bacteroids and the free-living USDA110 cells cultivated in rich medium (exponential growth phase in aerobic condition) were used as references (**Fig. 29a**).

Prior to quantification of transcript abundance or identification and quantification of protein accumulation, the transcriptome dataset was used to re-annotate the USDA110 genome with the EugenePP pipeline (23). This allowed the definition of 876 new CDS, ranging from 92 to 1091bp (median size = 215bp or 71.6 aa) with 11.5% of them having a predicted function or at least a match using InterProScan (IPR). This extends the total number of CDS in the USDA110 genome to 9171. Moreover, we also identified 246 ncRNAs, ranging from 49 to 765 bp (median = 76 bp), which were not annotated before. Proteomic evidence could be found for 28 new CDS (3.2% of the new CDS, median size = 97.6 aa). The complete reannotation of the genome is described in Table S1.

In the proteome dataset, 1808 USDA110 proteins were identified. Principal component analysis (PCA) of all the replicate samples and sample types revealed their partitioning according to the tested conditions. The first axis of the PCA (40.2% of the observed variance) separates bacteroid profiles from the exponential culture, whereas the second axis separates *G. max* bacteroids from *A. afraspera* bacteroids (14.9% of the observed variance; **Fig. 29b**). The samples of the transcriptome datasets are similarly distributed on the PCA plot, with a first axis explaining 42.6% of the observed variance and a second axis explaining 23.5% of the observed variance (**Fig. 29b**).

Although differences are less pronounced in the proteome dataset than in the transcriptome dataset, COG analysis shows similar profiles across functional categories, except for membrane proteins that are less well identified in proteomics than transcriptomics (**Supp.**

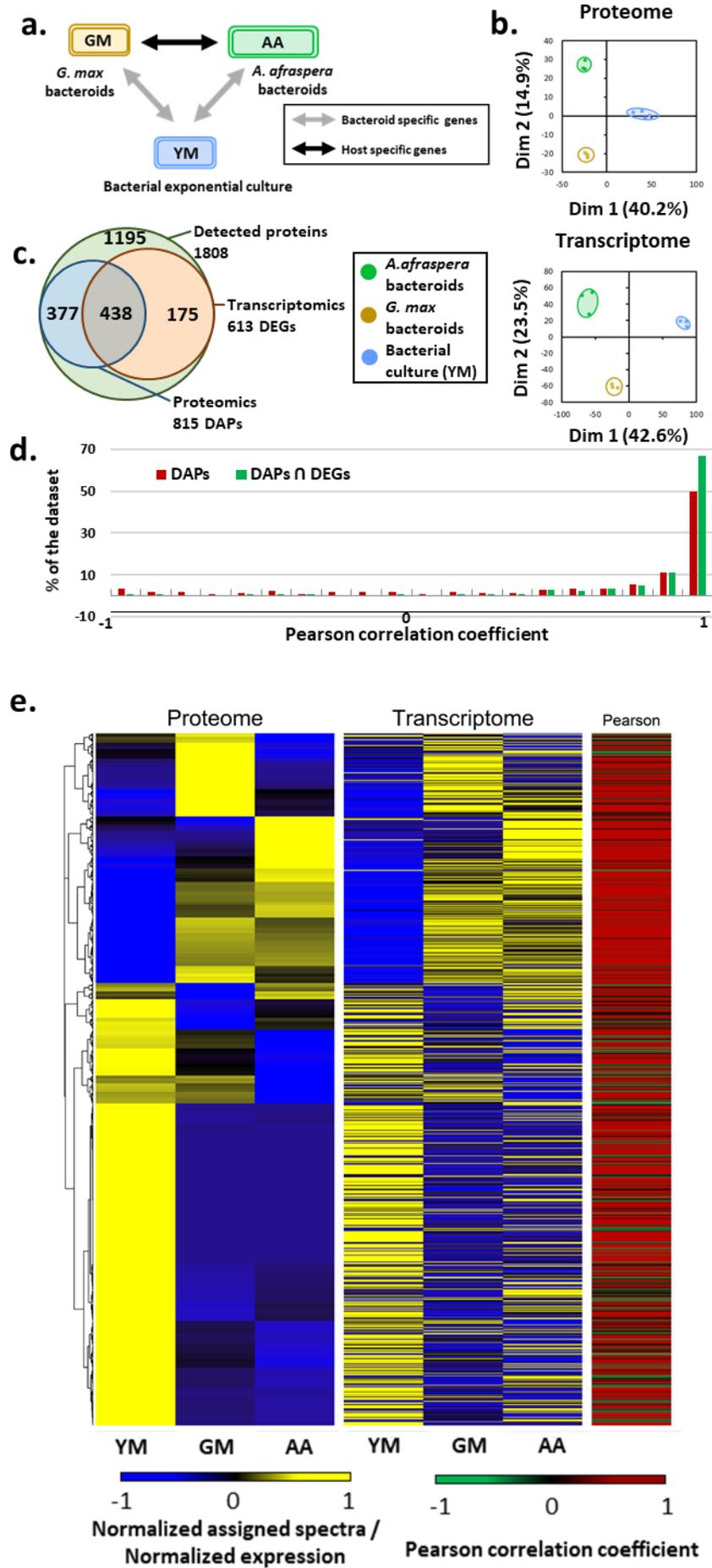


Fig. A4). In the transcriptomic dataset, 3150 genes are differentially expressed in at least one condition (differentially expressed genes or DEGs). Among the 1808 proteins identified, 815 show differential accumulation (differentially accumulated proteins or DAPs) and 438 of the cognate genes are also differentially expressed in the transcriptome datasets, whereas 175 DEGs are not DAPs (**Fig. 29c**).

We analyzed the Pearson correlation between transcriptomic and proteomic profiles and found that ~66% of the bacterial functions that show significant differences in both approaches display a high correlation coefficient ($r > 0.9$) whereas less than 1% of the functions show strong negative correlation ($r < -0.9$; **Fig. 29d**). This observation suggests that the transcriptome (which provides a more exhaustive view than the proteome) and the proteome provide a complementary picture of bacterial physiology, and they tend to show a congruent information if we restrict our analysis to the genes with differential accumulation/expression (**Fig. 29e**). However, there is still around 66% of the DEGs, which were detected by the proteomic analysis, that are not DAPs. Our description of the bacterial functions will be primarily based on the functions that are both DEGs and DAPs, as there is stronger evidence of their modulation in the tested conditions. The transcriptome alone will be used only when proteomics is not informative, for example to analyze regulons and stimulons that have been described previously in USDA110.

Symbiotic functions common to both types of USDA110 bacteroids

Among the 815 DAPs, 705 and 699 proteins are significantly differentially accumulated in *G. max* and *A. afraspera* respectively compared to the bacterial culture control. Strikingly, 646 proteins are commonly differentially accumulated in both plant nodules (Table S1).

In the transcriptomic dataset, 1999 DEGs, representing ~21% of the genome, were identified between the bacterial culture and the bacteroids, regardless of the host. Among them, 1076 genes displayed higher expression in nodules (including seven newly annotated ncRNAs and one newly annotated CDS among the 20 differentially expressed genes with highest fold change). An additional set of 923 genes were repressed *in planta* (including two

newly annotated ncRNAs and two newly annotated CDS among the 20 DEG with highest fold change, **Table S1**).

Restricting the analysis to the bacterial functions that are both differentially expressed (DEG) and differentially accumulated (DAP) *in planta* in both hosts as compared to the bacterial culture identified 222 genes/proteins, 150 being upregulated and 72 being repressed *in planta* respectively (**Fig. 30a**). Notably, six newly annotated genes are in this gene list including one putative regulator (*Bd110_01119*) that is induced during symbiosis. Among the functions commonly DEG and DAP *in planta*, only four functions showed opposite trends in proteomics and transcriptomics.

The proteome and transcriptome data provided a coherent view of the nitrogen fixation metabolism of *B. diazoefficiens* in the tested conditions. Key enzymes involved in microoxic respiration and nitrogen fixation were detected amongst the proteins having the highest spectra number in the nodule samples (**Fig. 30a, Table S1**) and the corresponding genes are among the most strongly expressed ones in bacteroids, while almost undetectable in the free-living condition. This includes for instance, the nitrogenase and the nitrogenase reductase subunits, which constitute the nitrogenase enzyme complex responsible for nitrogen conversion into ammonia. They belong to a locus of 21 genes from *blr1743* (*nifD*) to *bll1778* (*ahpC_2*), including the genes involved in nitrogenase cofactor biosynthesis, in electron transport to nitrogenase, and in microaerobic respiration, that are among the highest expressed ones in bacteroids of both host plants, both at the gene and protein expression level. The slightly higher level of the dinitrogenase reductase NifH detected in proteomics was not supported by western blot analysis, which showed apparent similar protein level in both bacteroid conditions (**Supp. Fig. A5**). Strikingly, the two bacteroid types did not show a notable difference in the expression of these genes and proteins, suggesting that the activation of the nitrogen fixation machinery is not a limiting factor underlying the suboptimal efficiency of strain USDA110 in *A. afraspera* nodules.

In addition to these expected bacteroid functions, many other proteins were identified that specifically and strongly accumulated in both nodule types. This is the case of the chaperonins GroEL1/GroES1, which are strongly upregulated and reach high gene expression and protein levels in both bacteroids. The upregulation of these chaperonins is remarkable

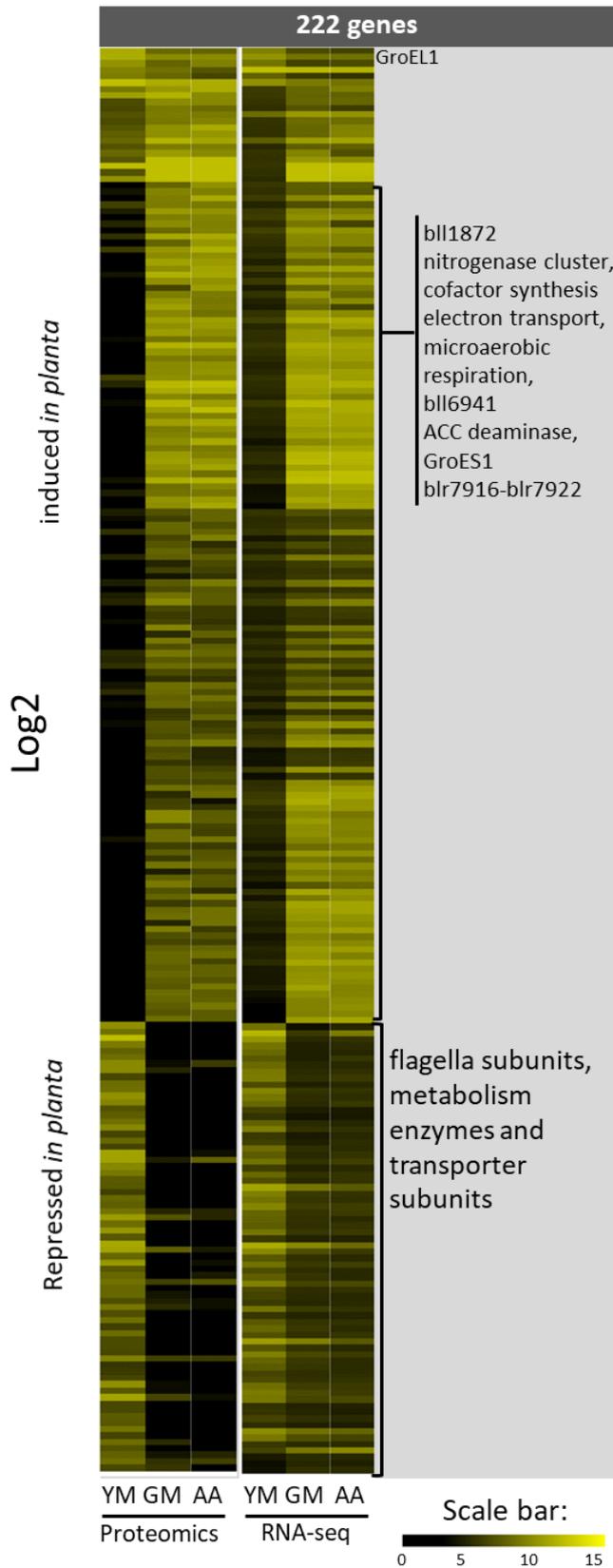
> **Figure 30 | Symbiosis and host-specific functions that display congruency between transcriptomics and proteomics a** | Heatmap with SOM clustering displaying bacterial functions that are commonly DAP and DEG *in planta* in both host plants as compared to the culture reference. **b** | Heatmap displaying bacterial functions that are commonly DEG and DAP in one host as compared to the other (upper panel: *A. afraspera* > *G. max*; lower panel: *G. max* > *A. afraspera*). In panels **a-b**, data are presented as log₂ of DESeq2 normalized read counts (RNA-seq) or spectral counting (Proteomics). YM: Yeast-Mannitol culture, GM: *G. max* bacteroids, AA: *A. afraspera* bacteroids.

because other GroEL/GroES (4, 5 and 7) proteins are also very strongly accumulated in a constitutive manner. This indicates that bacteroids have a high demand for protein folding, possibly requiring specific GroEL isoforms, a situation reminding the requirement of one out of five GroEL isoforms for symbiosis in *Sinorhizobium meliloti*, the symbiont of *Medicago sativa* (12,24). Another example of a bacteroid-specific function is the hydrogenase uptake system, whose gene expression was induced in both bacteroid types from nearly no expression in culture. Hydrogenase subunit *HupL_2* (*bll6941*) was found amongst the proteins displaying the highest spectra number in the nodule samples suggesting important electron recycling in bacteroids of the two hosts. Another one is the 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase (*blr0241*), which was also amongst the most strongly accumulated proteins in nodules and was significantly less abundant in free-living USDA110. An outer membrane protein (*bll1872*) belonging to the NifA regulon²⁵ was also strongly induced *in planta*, with a transcript level among the top 10 genes in *A. afraspera*. Additionally, a locus of seven genes (*blr7916-blr7922*) encoding an amidase enzyme and a putative peptide transporter composed of two transmembrane domain proteins, two ATPases and two solute-binding proteins was strongly upregulated in the two bacteroid types, with three protein being also over-accumulated *in planta* (**Fig. 30a; Table S1**).

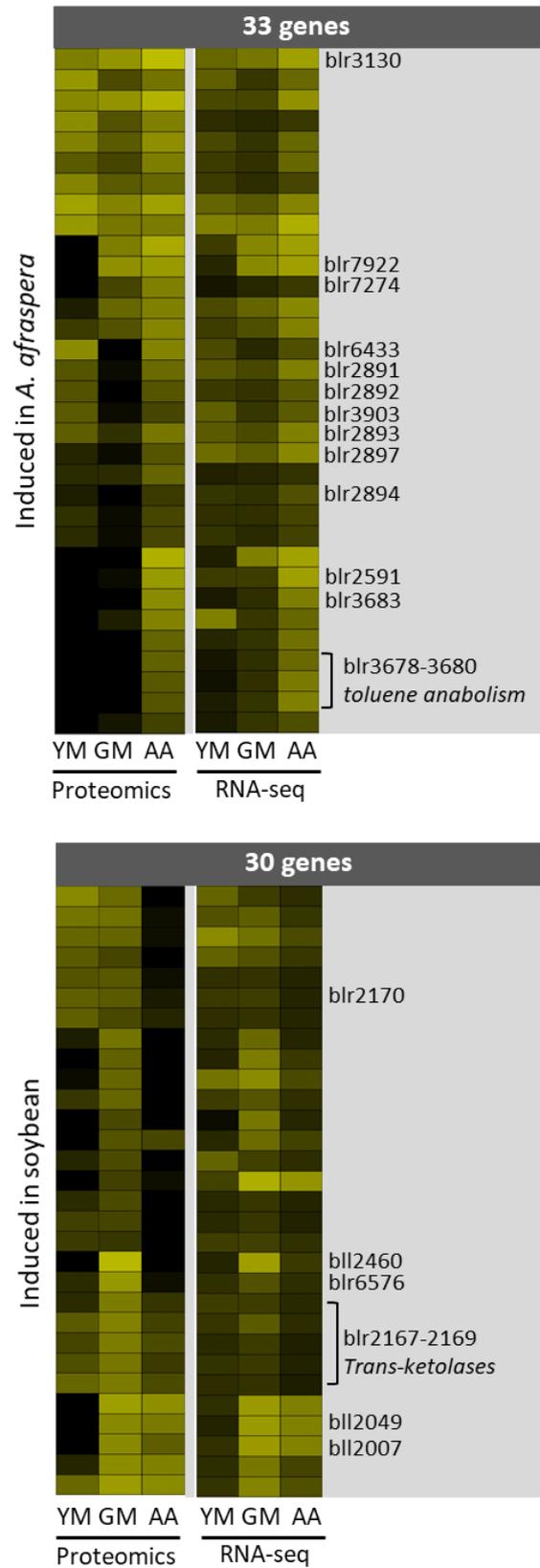
Oppositely, motility genes encoding flagella subunits (*bll5844-bll5846*), metabolic enzymes and transporter subunits are strongly downregulated during symbiosis and hardly detectable at the protein level *in planta* (**Fig. 30a**).

Taken together, these data show that both bacteroid types display a typical nitrogen fixation-oriented metabolism, with a partial shutdown of housekeeping functions. This indicates that despite the apparent reduced symbiotic efficiency of USDA110 in *A. afraspera* nodules, the bacterium fully expresses its symbiotic program within this non-native host as it does in soybean, its original host. Thus, the sub-optimal functioning of the *A. afraspera* nodules

a.



b.



does not seem to come from a bacterial defect to express the symbiotic program, but possibly from an unfavorable host microenvironment or from a lack of metabolic integration of these maladapted partners.

Host-specific functions

Comparison of the *A. afraspera* and *G. max* bacteroids revealed also significant differences in the proteomes and transcriptomes. At the transcriptomic level, 935 DEGs could be identified between the two bacteroid types (509 *A. afraspera* > *G. max* and 426 *G. max* > *A. afraspera*). One notable feature of the transcriptome is the identification of four newly annotated ncRNA and one new CDS amongst the 20 most induced DEGs in *A. afraspera* nodules and the presence of five newly annotated CDS amongst the 20 most induced DEGs in *G. max* nodules (**Table S1**). However, when considering only the functions that display congruent and significant differences in terms of transcripts and protein levels between plant hosts, we fall down to 63 genes/proteins, 33 being induced in *A. afraspera* nodules and 30 being induced in *G. max* nodules (**Fig. 30b**).

Interestingly, the phenylacetic acid degradation pathway (PaaABCDEIK, *blr2891-blr2897*) was highly expressed in *A. afraspera* nodules (although only PaaABCD and PaaK have been detected by proteomics), as well as a yet uncharacterized cluster of genes putatively involved in toluene degradation (*blr3675-blr3680*). The chaperone GroEL2 is also specifically induced in *A. afraspera*. Similarly, three S1 peptidases (Dop: *blr2591*, *blr3130* and *blr7274*) are highly expressed in the nodules of this latter host together with a RND efflux pump (*bll3903*) and a LTXXQ motif protein (*bll6433*), a motif also found in the periplasmic stress response CpxP26. The over-accumulation of these proteins suggests that bacteroids are facing stressful conditions during this interaction with *A. afraspera*. An uncharacterized ABC transporter solute binding protein (*blr7922*) was also overexpressed in *A. afraspera*.

One $\alpha\beta$ hydrolase (*blr6576*) and a TonB-dependent receptor-like protein (*bll2460*) were over-accumulated in a *G. max*-specific manner. Similarly, an uncharacterized metabolic cluster including transketolases (*blr2167-blr2170*), the heme biosynthetic enzyme HemN1 (*bll2007*)

and to a lesser extent an anthranilate phosphoribosyl-transferase (TrpD encoded by *bll2049*) are overexpressed in soybean nodules.

USDA110 transcriptomics data in the perspective of previously described regulons and stimulons

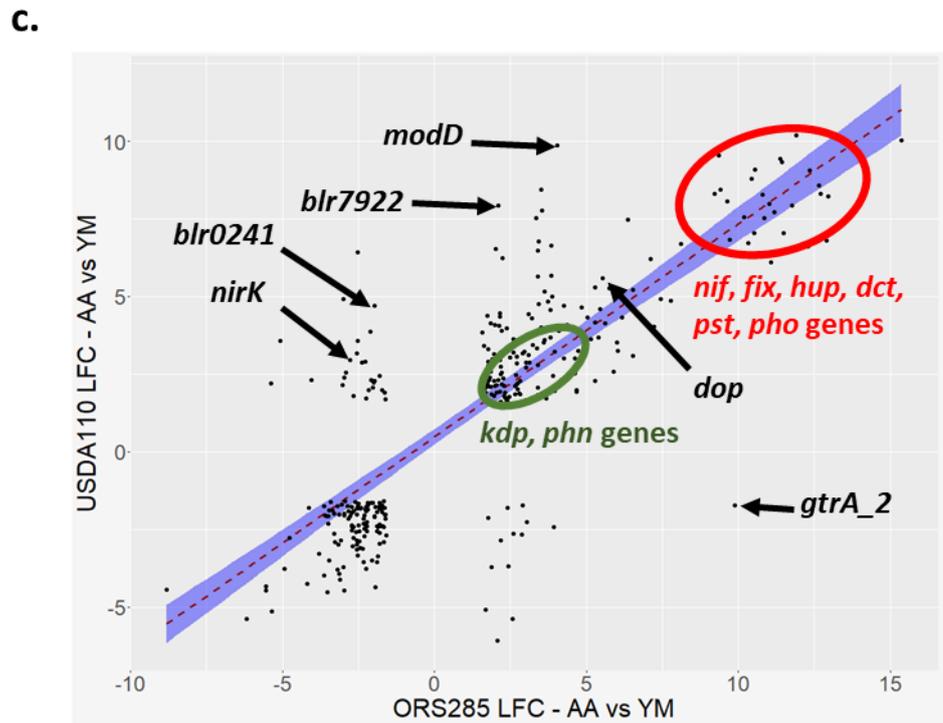
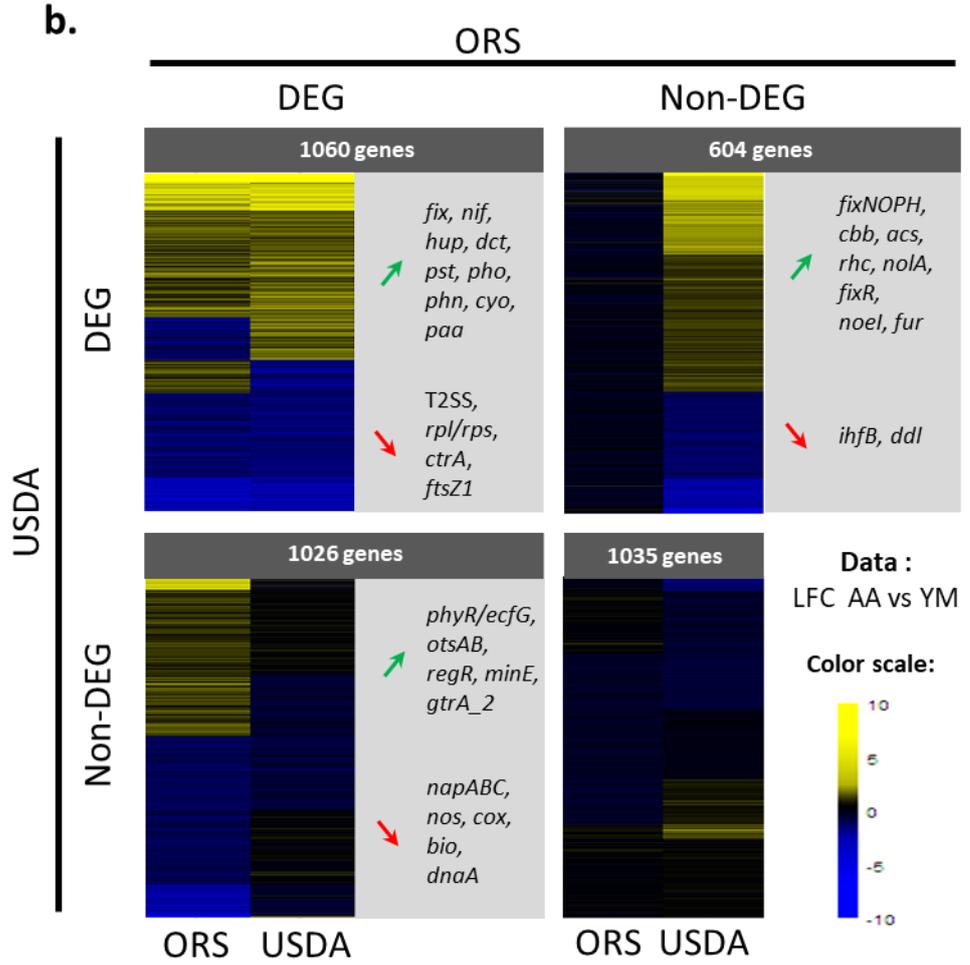
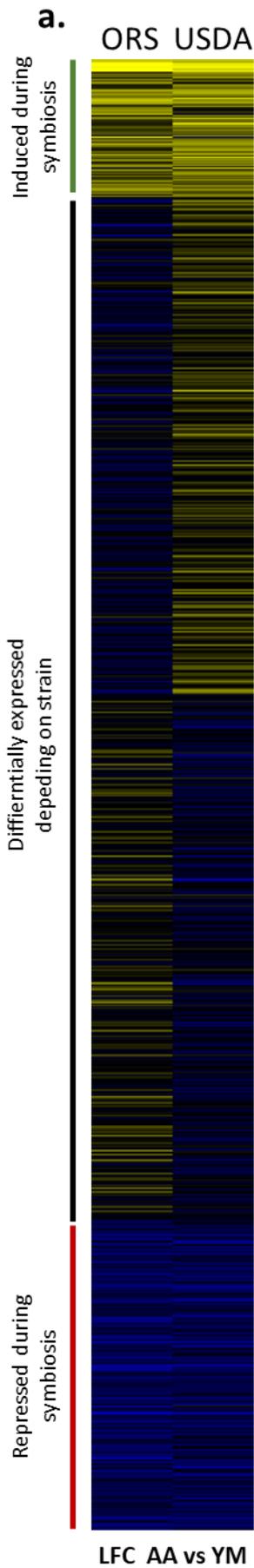
USDA110 is one of the best-characterized rhizobial strains in terms of transcriptomic responses to various stimuli as well as the definition of regulons (27). We analyzed the behavior of these previously defined gene networks in USDA110 in our dataset (**Table S2**). To initiate the molecular dialog that leads to nodule formation, plants secrete flavonoids like genistein in their root exudates, which are perceived by the rhizobia and trigger Nod factor production. At 14 dpi, when the nodule is formed and functioning, the genistein stimulon, which comprises the NodD1, NodVW, TtsI and LafR regulons, is not anymore activated in bacteroids. The symbiotic regulons controlled by NifA, FixK1, FixK2, FixLJ and sigma54 (RpoN) were activated *in planta*, indicating that nitrogen fixation was going on in both hosts. Accordingly, the nitrogen metabolism genes controlled by NtrC were activated *in planta*. Additionally, the PhyR/EcfG regulon involved in general stress response is not activated in bacteroids. Differences between hosts were however not observed for any of these regulons/stimulons. The only stimulon that showed differential expression between hosts is the one involved in aromatic compound degradation, which was highly expressed in *A. afraspera* nodules. Similar upregulation of the vanillate degradation pathway was observed in the transcriptome of *Bradyrhizobium* sp. ORS285 in *A. afraspera* and *A. indica* nodules (17), suggesting that Dalbergioid hosts display a higher aromatic compound content in nodules than *G. max*. In line with this hypothesis, some of the most differentially accumulated sets of proteins (*A. afraspera* > *G. max*) are involved in the degradation of phenylacetic acid (PaaABCDK and *bll0339*) suggesting that the bacterium converts phenylalanine (or other aromatic compounds) ultimately to fumarate through this route (**Fig. 30b**) (28). Similarly, enzymes of another pathway involved in phenolic compound degradation (*blr3675-blr3680*) are accumulated in *A. afraspera* nodules (**Fig. 30b, Table S1**).

> **Figure 31 | Expression pattern of *B. diazoefficiens* USDA110 and *Bradyrhizobium* sp. ORS285 orthologous genes *in planta* and in culture.** **a** | Heatmap after SOM clustering of all the orthologous genes of USDA110 and ORS285 obtained with Phyloprofile. Values present the *in planta* LFC calculated after the read counts of the culture control versus *A. afraspera* 14 dpi nodules. **b** | Heatmaps of the orthologous genes after filtering on the FDR (< 0.01) values. Selected genes are highlighted for each class of interest. **c** | Dot plot of the orthologous genes that are DEG (FDR < 0.01 and $|LFC| > 1.58$) *in planta* (i.e. in *A. afraspera* nodules) in both strains. The red dashed line is for the linear regression and the blue envelope shows a 0.95 confidence interval of the linear regression.

Expression pattern of orthologous genes between ORS285 and USDA110 in A. afraspera nodules

In a previous study (17), a transcriptome analysis was performed on *Bradyrhizobium* sp. ORS285 in interaction with *A. afraspera* and in culture. *Bradyrhizobium* sp. ORS285 is a strain that co-evolved with *A. afraspera*, leading to an efficient symbiosis hallmarked by TBD, *id est* cell elongation and polyploidization of the bacteroids. In order to compare gene expression of these two nodule-forming rhizobia in culture and *in planta*, we determined the set of orthologous genes between the two strains using the Phyloprofile tool of Mage Microscope website. This analysis yielded a total of 3725 genes (**Table S3**). The heatmap on **Figure 31a** presents the modulation of gene expression (LFC) between *A. afraspera* nodules and the bacterial culture for the orthologous genes in each bacterium, regardless of their statistical significance. When taking FDR < 0.01 in account, we identified sets of genes that are differentially expressed *in planta* in either bacterium or in both (**Fig. 31b**).

Only 343 genes displayed differential expression (FDR < 0.01 and $|LFC| > 1.58$) *in planta* in both bacteria as compared to their respective culture control (**Fig. 31c**). A majority of these genes (86.8%) exhibited congruent expression patterns. First, the *nif*, *fix* and *hup* genes are commonly and highly induced in both strains during their symbiotic life with *A. afraspera*, a hallmark of a functional symbiosis. However, there are differences in their expression level, with a higher expression of the symbiotic genes in ORS285 (*nifHDK* represent 12.5% of all reads in *A. afraspera* nodules) (17) than in USDA110 (*nifHDK* represent only 2.5% of all reads in *A. afraspera* nodules), consistently with a more efficient interaction occurring between ORS285 and *A. afraspera*. Additionally, the Kdp high affinity transport system, the phosphate (*pstCAB*, *phoU*, *phoE*, *phoC*) and phosphonate metabolism (*phnHIJKL*) are activated *in planta* in both



bacteria (**Fig. 31b-c**). The stress-marker dop protease gene is also induced in both bacteria in *A. afraspera* nodules (**Fig. 31c**).

Additionally, 1026 genes were differentially expressed solely in ORS285, and similarly there was 604 DEG specific to USDA110 (**Fig. 31b**). For example, the general secretory pathway seems to be specifically induced in ORS285. Oppositely, USDA110 displays an induction of the *rhcJQRU* genes which are involved in the injection of type three effector proteins that can be important for the establishment of the symbiosis whereas they are not induced or even

repressed in ORS285 (**Fig. 31b**). This is also the case of the nitrite reductase encoding gene *nirK* (*blr7089/BRAD285_v2_0763*; **Fig. 31c**). In addition, USDA110 induces the expression of an ACC deaminase (*blr0241*), while its ortholog is repressed in ORS285 (*BRAD285_v2_3570*) during symbiosis (**Fig. 31c**). Bacterial ACC deaminases can degrade ACC, a precursor of ethylene, and thereby modulate ethylene levels in the plant host and promote the nodulation process (29).

B. diazoefficiens USDA110 bacteroids undergo bona fide TBD in *A. afraspera* nodules despite very weak morphological and ploidy modifications

In a previous description of the *A. afraspera* - *B. diazoefficiens* USDA110 interaction, the typical TBD features were not observed and the bacteroids were very similar to those in *G. max* where no TBD occurs (16). At the molecular level, accumulation of the replication initiation factor DnaA is higher in soybean than in *A. afraspera* (**Table S1**). Similarly, the MurA peptidoglycan synthesis enzyme (encoded by *bl10822*) that may play a role in cell elongation during TBD was detected to similar levels in both bacteroids (**Table S1**). Taken together, the molecular data do not clearly indicate whether USDA110 bacteroids undergo TBD in *A. afraspera*. Therefore, we investigated the features of the USDA110 bacteroids in *A. afraspera* nodules in more detail.

We analyzed bacteroid differentiation features in USDA110 bacteroids extracted from soybean and *A. afraspera* nodules. The interaction between *A. afraspera* and *Bradyrhizobium* sp. ORS285 was used as a positive control for TBD features (9,30,31). TBD is characterized by cell elongation. We quantified cell length, width, area and shape of purified bacteroids and culture controls. Whereas ORS285 bacteroids were enlarged within *A. afraspera* nodules as

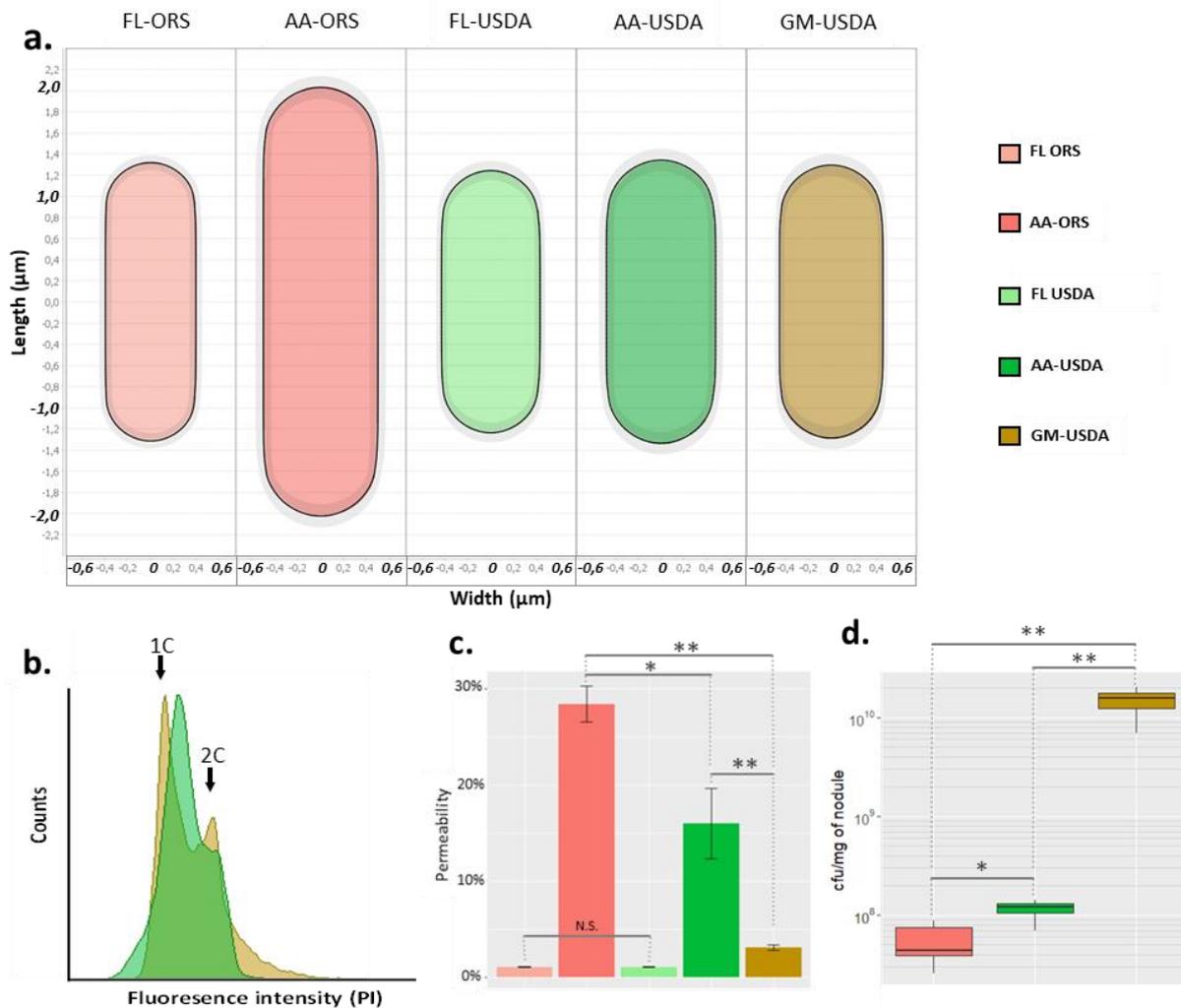


Figure 32 | *B. diazoefficiens* USDA110 displays atypical bacteroid differentiation features in *A. afraspera* nodules. a | Average cell shape of free-living bacteria and bacteroids determined by MicrobeJ (900 < n < 21 000). b | DNA content of USDA110 bacteroids extracted from soybean and *A. afraspera* determined by flow cytometry. c | Assessment of the permeability of USDA110 and ORS285 free-living cells and bacteroids 20 min after PI treatment. *: wilcoxon test, p-value < 0.01. Five biological replicates were performed for bacteroids and two for free-living bacteria. d | Viability of soybean and *A. afraspera* extracted bacteroids at 14 dpi. Asterisks point out significant differences according to a wilcoxon test. *: p-value < 0.05; **: p-value < 0.01. Data are representative of 10 independent plants.

compared to their free-living counterparts, USDA110 bacteroids were similar to free-living bacteria in both soybean and *A. afraspera* (Fig. 32a; Supp. Fig. A6). Another feature of TBD is endoreduplication. Analysis of the bacterial DNA content of ORS285 bacteroids in *A. afraspera* by flow cytometry shows peaks at 6C and more (9). As expected, USDA110 bacteroids in *G. max* yields only two peaks, at 1C and 2C, similarly to the cycling cells in the bacterial culture sample (Fig. 32b) (16). Strikingly, similar results were obtained for USDA110 in *A. afraspera*. Thus, with respect to the DNA content and cell size, the USDA110 bacteroids do not display

the typical TBD features in *A. afraspera* nodules. Loss of membrane integrity is a third hallmark of TBD that likely strongly contributes to the loss of viability of bacteroids. Time-course analysis of propidium iodide (PI) uptake by bacteroids and the corresponding culture controls were performed to assess bacteroid permeability (**Supp. Fig. A7**). Twenty minutes after PI application, USDA110 bacteroids from *A. afraspera* display an increased permeability that is much closer to ORS285 bacteroids in interaction with *A. afraspera* than to the low permeability of USDA110 bacteroids from *G. max* nodules (**Fig. 32c**). Also, the free-living counterparts exhibit a very low permeability. Taken together, this suggests that the envelope of USDA110 bacteroids is more permeable in the NCR-producing *A. afraspera* nodules, even if it does not reach the permeability level of the ORS285 strain. To analyze bacterial viability, bacteroids extracted from nodules were plated and the colony forming units (cfu) were determined (**Fig. 32d**). In *G. max*, USDA110 formed 1.46×10^{10} colonies/mg nodule (~100% survival). Oppositely, ORS285 formed only 5.42×10^7 colonies/mg nodule in *A. afraspera* (~0.5% survival). Interestingly, USDA110 formed 1.13×10^8 colonies/mg nodule in *A. afraspera* (~1% survival), indicating that, despite the absence of cell enlargement and endoreduplication USDA110 bacteroids lose their viability and undergo a bona fide terminal differentiation in *A. afraspera*. Thus, in the NCR-producing plant *A. afraspera*, USDA110 bacteroids display a disconnection of the four canonical TBD features (ie. cell size, ploidy level, membrane permeability and cell viability).

vi. Discussion

A. afraspera triggers atypical but terminal differentiation of USDA110 bacteroids

In a previous study, we noticed that, in *A. afraspera*, USDA110 forms a functional symbiosis although bacteroids do not display features that are usually associated with TBD (16). Here we show that no endoreduplication and cell elongation of USDA110 occur in terminally differentiated bacteroids that fix nitrogen in a suboptimal way. Accordingly, the protein level of DnaA, the genome replication initiator, was higher in soybean than in *A. afraspera* bacteroids and the MurA level was not different between bacteroid conditions, confirming that polyploidization and cell elongation did not occur in this host. Such unusual terminal bacteroid differentiation is reminiscent of the bacteroids in *Glycyrrhiza uralensis*. This plant of the IRLC

expresses NCR peptides (11). However, one of its compatible symbionts, *Sinorhizobium fredii* strain HH103, does not undergo any loss of viability, no change in DNA content and no cell elongation (32), while another symbiont, *Mesorhizobium tianshanense* strain HAMBI 3372 showed all TBD features (33). The influence of the bacterial genotype on terminal/non-terminal differentiation of bacteroids was also suggested in *Medicago truncatula* in which, the gene *hrrP* might confer to some *Sinorhizobium* strains a resistance against the differentiation process triggered by some *M. truncatula* ecotypes (34). In these two IRLC plants (*ie. M. truncatula* and *G. uralensis*), bacteria undergo a complete TBD or no TBD at all in a strain-dependent manner, but there is no clear uncoupling of the features of TBD (cell elongation/endoreduplication/altering viability) as shown here in the case of *B. diazoefficiens* USDA110 – *A. afraspera*.

The surprising differentiation of USDA110 in *A. afraspera* nodules raises questions about the molecular mechanisms supporting this phenomenon. We consider two possible hypotheses: strain USDA110 might be more sensitive to the differentiation factors of the host than strain ORS285 and be rapidly “terminally” differentiated, before the other differentiation features, that are potentially important for symbiotic efficiency, can take place. Alternatively, USDA110 might be resistant to the plant effectors that trigger the elongation and polyploidization features.

In agreement with the latter possibility, the application of NCR peptides has very limited effect on strain USDA110 as compared to *S. meliloti* and to other plant-associated bacteria (16,35). NCR insensitivity may be due to the thick hopanoid layer that is present in the outer membrane of strain USDA110, as the hopanoid biosynthesis mutant *hpnH* is more sensitive to NCR peptides and shows symbiotic defects in *A. afraspera* but not in *G. max* (36). Moreover, the altered peptidoglycan structure in the strain USDA110 DD-carboxypeptidase mutant resulted in an increased TBD process with endoreduplicated and elongated bacteroids in *A. afraspera* (16). This suggests that the envelope of strain USDA110 prevents a canonical TBD to occur. Possibly, NCR peptides are not able to reach their intracellular targets required to induce endoreduplication and cell division arrest, while their effect on cell viability through pore formation and membrane destabilization is still effective.

A survey of TBD in the legumes has identified multiple occurrences of the process in several subclades of the legumes but found that the majority of legumes do not have TBD (37). The classification in this study was based on a morphological analysis of the bacteroids. Ancestral state reconstruction based on this classification suggested that the non-differentiated bacteroids are ancestral and that TBD evolved at least five times independently in legumes (37). The discovery of bacteroids that are terminally differentiated without any obvious morphological changes opens the possibility that the occurrence of TBD might be underestimated in the legume family. Similarly, in the IRLC clade, the extent of morphological bacteroid differentiation was correlated to the size of the cationic NCR peptides repertoire and in legumes with few NCR peptides, the morphological modification of bacteroids can be minor (11,33). In addition, at the molecular level, TBD is originally ascribed to the production of symbiotic antimicrobial peptides, the NCRs, by nodules (7), but more recently, other types of antimicrobial peptides such as the NCR-like, GRP, MBP1 and CAPE peptides specifically produced in nodules of different plants were proposed to contribute to bacteroid differentiation (9,38-40). Thus, if TBD would indeed be more widespread than currently estimated on the basis of morphological bacteroid features, the currently proposed evolutionary scenario of bacteroid formation might require revision.

Terminal differentiation is associated with specific stress response

The TBD of strain USDA110 in *A. afraspera* is associated with a higher accumulation of stress markers compared to the *G. max* bacteroids. These markers include four proteases (Dop, Lon_2, *blr3130* and *blr7274*) and one chaperonin (GroEL_2). Similar induction of proteases and chaperonins have been reported in NCR-treated *S. meliloti* cultures (35), indicating that this response may be linked to the perception of *A. afraspera* NCR-like peptides in USDA110.

The genes encoding these stress related proteins are not part of the well-characterized general stress response (GSR) controlled by the PhyR/EcfG signaling cascade in *B. diazoefficiens* USDA110 (41). On the other hand, we found that the PhyR/EcfG regulon in USDA110 is not activated in the bacteroids of both host plants (**Table S2**). This observation contrasts with our previous study of *Bradyrhizobium* sp. ORS285 transcriptome during symbiosis with *Aeschynomene* plants, which showed that the PhyR/EcfG cascade was upregulated *in planta* (17). Nevertheless, the expression of the Dop protease was induced in *A. afraspera* in both

bacteria (**Fig. 31c**). Together, the omics data suggest that bacteroids of *Bradyrhizobium* spp. activate stress-related genes in the TBD-inducing *A. afraspera* host, but differences exist in the activation of specific stress responses at the strain level.

Correlation between bacteroid differentiation features and symbiotic efficiency for the plant

TBD is associated with the massive production of symbiotic antimicrobial peptides such as NCR, NCR-like and CAPE peptides in different plants (5,9,38,40). They represent ~10% of the nodule transcriptomes in *M. truncatula* (analysis of the data from ref 42) and their production is thus potentially a strong energetic cost for the plant, raising questions about the benefits of the TBD process. TBD appeared independently in different legume clades (9,37), suggesting that plants imposing this process obtain an advantage which might be a higher symbiotic benefit. Increased symbiotic efficiency has indeed been observed in hosts imposing TBD (17,43,44). The findings reported here, comparing bacteroids and symbiotic efficiency in *A. afraspera* infected with strain ORS285 and strain USDA110, are in agreement with this hypothesis. Also, in the symbiosis of *M. truncatula* in interaction with different *S. meliloti* strains, a similar correlation was observed between the level of bacteroid differentiation and the plant growth stimulation (45). However, the simultaneous analysis of the bacteroid differentiation and symbiotic performance of an extended set of *Aeschynomene* – *Bradyrhizobium* interactions has shown that, perhaps not unexpectedly, the symbiotic efficiency of the plant-bacterium couple is not solely correlated with bacteroid differentiation and that other factors can interfere with the symbiotic efficiency as well (46).

vii. Conclusion

Bradyrhizobium diazoefficiens USDA110 is a major model in the legume-rhizobium symbiosis, mainly thanks to its interaction with *G. max*, the worldwide most cultivated legume. Although omic studies have been conducted in this strain in symbiosis with various hosts (13,25), this is the first time that this bacterium is studied at the molecular level in symbiosis with a NCR-producing plant that normally trigger a typical terminal bacteroid differentiation in its symbionts. The symbiosis between USDA110 and *A. afraspera* is functional even if nitrogen fixation and plant benefits are sub-optimal.

Terminal bacteroid differentiation is taking place in the NCR-producing host *A. afraspera*, as bacterial viability is impaired in USDA110 bacteroids, whereas morphological changes and the cell cycle switch to endoreduplication are not observed. We also show by combining proteomics and transcriptomics that the bacterial symbiotic program is expressed in *A. afraspera* nodules in a similar way as in *G. max*, although host-specific patterns were also identified. However, the bacterium is under stressful conditions in the *A. afraspera* host, possibly due to the production of NCR-like peptides in this plant. Integration of datasets from different bacteria in symbiosis with a single host, like ORS285 and USDA110 in symbiosis with *A. afraspera*, shed light on the differences in the stress responses activated in *A. afraspera* and confirmed that the symbiosis is functional but suboptimal in this interaction. The molecular data presented here provide a set of candidate functions that could be analyzed for their involvement in the adaptation to a new host and to the TBD process.

viii. Material and Methods

Bacterial cultures and bacteroid extraction

B. diazoefficiens USDA11047 and *Bradyrhizobium* sp. ORS285 were cultivated in YM medium at 30°C in a rotary shaker (48). For transcriptomic analysis, culture samples (OD_{600nm} = 0.5) were collected and treated as in Chapelle et al. (2015) (49).

G. max ecotype Williams 82 and *A. afraspera* seeds were surface-sterilized and the plants were cultivated and infected with rhizobia for nodule formation as described in Barrière et al. (2017) 16. Nodules were collected at 14 days post inoculation (dpi), immediately immersed in liquid nitrogen and stored at -80°C until use. Each tested condition (in culture and *in planta*) was produced in biological triplicates.

Phylogeny analysis

Nucleotide sequence of *matK* genes were collected on NCBI using accession numbers described in references 50 and 51 and analyzed on phylogeny.fr (www.phylogeny.fr). They were aligned using ClustalW with manual corrections, before running a phyML (GTR - Gamma model) analysis with 500 bootstraps. A Bayesian inference tree was also generated (GTR + G +

I) and provided similar topology as the maximum likelihood tree (data not shown). Trees were visualized and customized using TreeDyn.

Genome annotation and RNA-seq analysis

Nodule and bacterial culture total RNA was extracted and treated as previously described in 17. Oriented (strand-specific) libraries were produced using the SOLiD Total RNA-seq kit (Life Technologies) and were sequenced on a SOLiD 3 station yielding ~40 million 50bp single reads. Trimming and normalization of the reads were performed using the CLC workbench software. Subsequently, the reads were used to annotate the genome using EugenePP (23), and the mapping was performed using this new genome annotation. Analysis of the transcriptome using DE-seq2 and data representation were performed as previously described (17). Differentially expressed genes (DEG) showed an absolute log₂ fold change (|LFC|) > 1.58 (ie. fold change > 3) with a false discovery rate (FDR) < 0.01.

Proteomic analysis

Bacteroids were extracted from 14 dpi frozen nodules (6), while bacterial culture samples were collected as above, and the bacterial pellets were resuspended in -20°C acetone and lysed by sonication. Protein solubilization, dosage, digestion (trypsin 2% w/w) and solid phase extraction (using Phenomenex polymeric C18 column) were performed as described before (52). Peptides from 800 ng of proteins were analyzed by LC-MS/MS with a Q Exactive mass spectrometer (Thermo Electron) coupled to a nanoLC Ultra 2D (Eksigent) using a nanoelectrospray interface (non-coated capillary probe, 10 µ i.d.; New Objective). Peptides were loaded on a Biosphere C18 trap-column (particle size: 5 µm, pore size: 12 nm, inner/outer diameters: 360/100 µm, length: 20 mm; NanoSeparations) and rinsed for 3 min at 7,5µl minute of 2% Acetonitrile (ACN), 0,1% Formic acid (FA) in water. Peptides were then separated on a Biosphere C18 column (particle size: 3 µm, pore size: 12 nm, inner/outer diameters: 360/75 µm, length: 300 mm; NanoSeparations) with a linear gradient from 5% of 0,1% FA in ACN (buffer B) and 95% of 0,1% FA in Water (buffer A) to 35% of buffer B and 65% of buffer A in 80 min at 300nl/min followed by a rinsing step at 95% of buffer B and 5% of buffer A for 6 min and a regeneration step with parameters of the start of the gradient for 8 min. peptide ions were analyzed using Xcalibur 2.1 software in data dependent mode with the following parameters:

(I) full ms was acquire for the 400-1400 m/z range at a resolution of 70000 with an AGC target of $3 \cdot 10^6$; (ii) MS² scan was acquired at a resolution of 17500 with an agc target of $5 \cdot 10^4$, a maximum injection time of 120 ms and an isolation window of 3 m/z. The normalized collision energy was set to 27. MS² scan was performed for the eight most intense ions in previous full MS scan with an intensity threshold of $1 \cdot 10^3$ and a charge between 2 and 4. Dynamic exclusion was set to 50s. After conversion to mzXML format using msconvert (version 3.0.3706) (53), data were searched using X!tandem (version 2015.04.01.1) (54) against the USDA110 reannotated protein database and a homemade database containing current contaminants. In a first pass trypsin was set to strict mode and cysteine carbamidomethylation as a fixed modification and methionine oxidation, protein N-terminal acetylation with or without protein N-terminal methionine excision, N-terminal glutamine and carbamidomethylated cysteine deamidation, N-terminal glutamic dehydration as potential modifications. In a refine pass, semi-enzymatic peptides were allowed. Proteins inference was performed using X!TandemPipeline (version 3,4,3) (55). A protein was validated with an E-value $< 10^{-5}$ and 2 different peptides with an E-value < 0.05 . Protein from contaminant database (*Glycine max* proteins and unpublished *Aeschynomene* Expressed Sequence Tags) were removed after inference. Proteins were quantified using the spectral counting method (56). To discriminate differentially accumulated proteins (DAPs), ANOVA analysis was performed on the spectral counts and proteins were considered DAP when p-value < 0.05 .

Metabolomic analysis

Metabolites and cofactors were extracted from lyophilized nodules and analyzed by GC-MS and LC-MS respectively according to Su et al. (2016) (57).

Plant biomass and nitrogen fixation analysis

Dry mass of shoot, root and nodules was measured, and shoot-root mass ratio was calculated. The mass gain per g of dry nodule was calculated as the difference between total mean masses of the plants of interest and of the non-inoculated plants, divided by the mean mass of nodules. Thirty plants were used per condition. Nitrogenase activity was assessed by Acetylene Reduction Assay (ARA) on ten plants per condition as previously described (31). The elemental analysis of leaf carbon and nitrogen content was performed as described in reference 18.

Analysis of B. diazoefficiens USDA110 regulons and stimulons

Gene sets defined as regulons and stimulons were collected from the literature and the regulons/stimulons were considered as activated/repressed when $\geq 40\%$ of the corresponding genes were DEG in a host plant as compared to the culture condition.

Comparison of orthologous gene expression between B. diazoefficiens USDA110 and Bradyrhizobium sp. ORS285

The list of orthologous genes between USDA110 and ORS285 was determined using the Phyloprofile tool of the MicroScope-MAGE platform (58), with identity threshold of 60%, maxLrap > 0 and minLrap > 0.8. The RNA-seq data from reference 17 and those of this study were used to produce heatmaps, for the genes displaying FDR < 0.01 (*A. afraspera* vs. YM), using R (v3.6.3) and drawn using pheatmap (v1.0.12) coupled with kohonen (v3.0.10) for gene clustering using the Self Organizing Maps (SOM) method. The DEG in both organisms (*A. afraspera* vs YM) were plotted for USDA110 and ORS285.

Analysis of TBD features

Bacteroids were extracted from 14 dpi nodules and analyzed using a CytoFLEX S (Beckman-Coulter) (31). For ploidy and live/dead analyses, samples were stained with propidium iodide (PI, ThermoFisher, 50 $\mu\text{g}\cdot\text{mL}^{-1}$ final) and Syto9 (ThermoFisher, 1.67 μM final). PI permeability was assessed over time on live bacteria. *Bradyrhizobium* sp. ORS285.pMG103-nptII-GFP30 and *B. diazoefficiens* USDA110 sYFP2-159 strains were used to distinguish bacteroid from debris during flow cytometry analysis. For each time point, the suspension was diluted 50 times for measurement in the flow-cytometer. The percentage of bacteroids permeable to PI was estimated as the ratio of PI-positive over total bacteroids (GFP/YFP positive). Heat-killed bacteroids were used as positive control to identify the PI-stained bacteroid population.

For bacteroid viability assays, nodules were collected and surface sterilized (1 min NaClO 0.4%, 1 min 70% ethanol, two washes in sterile water). Bacteroids were subsequently prepared as previously described (31) and serially diluted and plated (five μl per spot) in triplicate on YM medium containing 50 $\mu\text{g}\cdot\text{mL}^{-1}$ carbenicillin. Colony-forming units (cfu) were counted five days post plating and divided by the total nodule mass.

Bacterial cell shape, length and width were determined using confocal microscopy image analysis. Bacteroid extracts and stationary phase bacteria cultures were stained with 2.5 nM Syto9 for 10 minutes at 37°C and mounted between slide and coverslip. Bacteria imaging was performed on a SP8 laser scanning confocal microscope (Leica microsystems) equipped with hybrid detectors and a 63x oil immersion objective (Plan Apo, NA: 1.4, Leica). For each condition, multiple z-stacks (2.7 μm width, 0.7 μm step) were automatically acquired (excitation: 488 nm; collection of fluorescence: 520–580 nm).

Prior to image processing, each stack was transformed as a maximum intensity projection using ImageJ software (<https://imagej.nih.gov/ij/>). Bacteria detection was performed with MicrobeJ (<https://www.microbej.com/>) (60). First, bacteria were automatically detected on every image using an intensity based thresholding method with a combination of morphological filters (area: 1–20 μm²; length: 1 μm–∞; width: 0.5–1.3 μm) and every object was fitted with a “Rod-shaped” bacteria model. To ensure high data quality every image was manually checked to remove false positive (mainly plant residues) and include rejected objects (mainly fused bacteria). Then the morphology measurements and figures were directly extracted from MicrobeJ. ORS285 in culture and in symbiosis with *A. afraspera* were used as references for the analysis of TBD features.

Western blot analysis

Detection of NifH by western blotting was performed using a commercial polyclonal antibody against a NifH peptide (Agrisera) respectively. The western blotting was carried out as previously described (61) using bacterial exponential (OD_{600nm} = 0.5) and stationary (OD_{600nm} > 2.5) phase cultures as well as 14 dpi nodule-extracted bacteroids.

ix. Acknowledgments

The authors would like to thank Dora Latinovics for production and sequencing of the RNA-seq libraries and Mélisande Blein-Nicolas for her advices regarding the statistical analysis of the proteomic dataset. Q.N. and F.L. were supported by a PhD fellowship from the “Université Paris-Sud”. The present work has benefited from the core facilities of Imagerie-Gif (<http://www.i2bc.paris-saclay.fr>), member of IBiSA (<http://www.ibisa.net>), supported by ‘France-BioImaging’ (ANR-10-INBS-04-01), and the Labex ‘Saclay Plant Sciences’ (ANR-11-IDEX-0003-

02). This work was funded by the 'Agence Nationale de la Recherche', grants n° ANR-13-BSV7-0013 and ANR-17-CE20-0011 and used resources from the National Office for Research, Development and Innovation of Hungary, grant n° 120120 to A.K.

x. Authors' contribution

QN, FL, PM, BGo and BA designed the work. QN, FL, AC, TB, MB, FGU, ES, ST and OP performed the experiments. QN, FL, MB, ES, YD, BGa, FGi, IN, AK, MZ, PM, BGo and BA analyzed the data. QN, FL, PM, BGo and BA wrote the paper.

xi. References

1. Erisman, J.W., Galloway, J.N., Seitzinger, S., Bleeker, A., Dise, N.B., Petrescu, A.M.R., et al. (2013) Consequences of human modification of the global nitrogen cycle. *Philos Trans R Soc B Biol Sci* 368: 20130116–20130116.
2. Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D.B., Huang, Y., et al. (2017) Temperature increase reduces global yields of major crops in four independent estimates. *Proc Natl Acad Sci U S A* 114: 9326–9331.
3. Oldroyd, G. (2013) Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat Rev Microbiol* 11: 252–263.
4. Gourion, B. and Alunni, B. (2018) Strain-specific symbiotic genes: A new level of control in the intracellular accommodation of rhizobia within legume nodule cells. *Mol Plant-Microbe Interact* 31: 287–288.
5. Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., et al. (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: 161–73.
6. Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., et al. (2006) Eukaryotic control on bacterial cell cycle and differentiation in the rhizobium-legume symbiosis. *Proc Natl Acad Sci U S A* 103: 5230–5235.
7. Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., et al. (2010) Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327: 1122–1126.
8. Guefrachi, I., Nagymihaly, M., Pislariu, C.I., Van de Velde, W., Ratet, P., Mars, M., et al. (2014) Extreme specificity of NCR gene expression in *Medicago truncatula*. *BMC Genomics* 15(1):712.
9. Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., et al. (2015) Convergent evolution of rnod symbiont differentiation in Dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by nodule-specific cysteine-rich peptides. *Plant Physiol* 169: 1254–1265.
10. Alunni, B., and Gourion, B. (2016) Terminal bacteroid differentiation in the legume-rhizobium symbiosis: nodule-specific cysteine-rich peptides and beyond. *New Phytol* 211: 411–417.
11. Montiel, J., Downie, J.A., Farkas, A., Bihari, P., Herczeg, R., Bálint, B., et al. (2017) Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proc Natl Acad Sci U S A* 114: 5041–5046.
12. Farkas, A., Maróti, G., Dürgo, H., Györgypal, Z., Lima, R.M., Folkl-Medzihradsky, K., Kereszt, A., Mergaert, P., and Kondorosi, E. (2014) *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proc Natl Acad Sci U S A* 111: 5183–5188.
13. Koch, M., Delmotte, N., Rehrauer, H., Vorholt, J.A., Pessi, G., and Hennecke, H. (2010) Rhizobial adaptation to hosts, a new facet in the legume root-nodule symbiosis. *Mol Plant-Microbe Interact* 23: 784–790.
14. Renier, A., Maillet, F., Fardoux, J., Poinot, V., Giraud, E., and Nouwen, N. (2011) Photosynthetic *Bradyrhizobium* Sp. strain ORS285 synthesizes 2-O-methylfucosylated lipochitooligosaccharides for nod gene-dependent interaction with *Aeschynomene* plants. *Mol Plant Microbe Interact* 24: 1440–7.

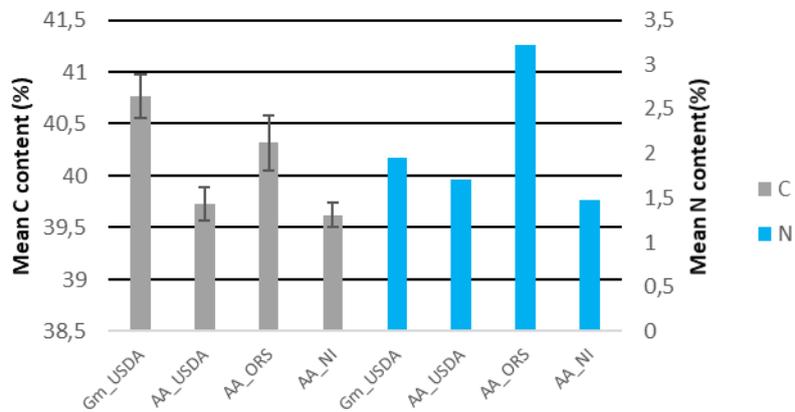
15. Ledermann, R., Bartsch, I., Müller, B., Wülser, J., and Fischer, H.M. (2018) A functional general stress response of *Bradyrhizobium diazoefficiens* is required for early stages of host plant infection. *Mol Plant-Microbe Interact* 31: 537–547.
16. Barrière, Q., Guefrachi, I., Gully, D., Lamouche, F., Pierre, O., Fardoux, J., et al. (2017) Integrated roles of BclA and DD-carboxypeptidase 1 in *Bradyrhizobium* differentiation within NCR-producing and NCR-lacking root nodules. *Sci Rep* 7: 1–13.
17. Lamouche, F., Gully, D., Chaumeret, A., Nouwen, N., Verly, C., Pierre, O., et al. (2019) Transcriptomic dissection of *Bradyrhizobium* sp. strain ORS285 in symbiosis with *Aeschynomene* spp. inducing different bacteroid morphotypes with contrasted symbiotic efficiency. *Environ Microbiol.* 21: 3244–3258.
18. Lamouche, F., Chaumeret, A., Guefrachi, I., Barrière, Q., Pierre, O., Guérard, F., et al. (2019) From intracellular bacteria to differentiated bacteroids: transcriptome and metabolome analysis in *Aeschynomene* nodules using the *Bradyrhizobium* sp. ORS285 bclA mutant. *J. Bacteriol* 201: e00191-19.
19. Andrews, M., Raven, J.A., Lea, P.J., and Sprent, J.I. (2006) A role for shoot protein in shoot-root dry matter allocation in higher plants. *Ann Bot* 97: 3-10.
20. Collier, R., and Tegeder, M. (2012) Soybean ureide transporters play a critical role in nodule development, function and nitrogen export. *Plant J.* 72: 355-67.
21. Ross, E.J.H.H., Kramer, S.B., and Dalton, D.A. (1999) Effectiveness of ascorbate and ascorbate peroxidase in promoting nitrogen fixation in model systems. *Phytochemistry* 52: 1203–1210.
22. Bashor, C.J. and Dalton, D.A. (1999) Effects of exogenous application and stem infusion of ascorbate on soybean (*Glycine max*) root nodules. *New Phytol* 142: 19–26.
23. Sallet, E., Gouzy, J., and Schiex, T. (2014) EuGene-PP: a next-generation automated annotation pipeline for prokaryotic genomes. *Bioinformatics* 30: 2659–61.
24. Bittner, A.N., Foltz, A., and Oke, V. (2007) Only one of five groEL genes is required for viability and successful symbiosis in *Sinorhizobium meliloti*. *J Bacteriol* 189: 1884–9.
25. Lardi, M., Murset, V., Fischer, H. M., Mesa, S., Ahrens, C. H., Zamboni, N. et al. (2016) Metabolomic profiling of *Bradyrhizobium diazoefficiens*-induced root nodules reveals both host plant-specific and developmental signatures. *Int J Mol Sci* 17(6) : 815.
26. Thede, G. L., Arthur, D. C., Edwards, R. A., Buelow, D. R., Wong, J. L., Raivio, T. L., & Glover, J. N. (2011) Structure of the periplasmic stress response protein CpxP. *Journal Bacteriol* 193(9): 2149–57.
27. diCenzo, G.C., Zamani, M., Checcucci, A., Fondi, M., Griffiths, J.S., Finan, T.M., and Mengoni, A. (2018) Multi-disciplinary approaches for studying rhizobium – legume symbioses. *Peer J* cjm-2018-0377.
28. Teufel, R., Mascaraque, V., Ismail, W., Voss, M., Perera, J., Eisenreich, W., et al. (2010) Bacterial phenylalanine and phenylacetate catabolic pathway revealed. *Proc Natl Acad Sci U S A* 107: 14390–5.
29. Okazaki, S., Nukui, N., Sugawara, M., and Minamisawa, K. (2004) Rhizobial strategies to enhance symbiotic interactions: rhizobitoxine and 1-aminocyclopropane-1-carboxylate deaminase. *Microb. Environ* 19: 99–111.
30. Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., et al. (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.
31. Guefrachi, I., Pierre, O., Timchenko, T., Alunni, B., Barrière, Q., Czernic, P., et al. (2015) *Bradyrhizobium* BclA is a peptide transporter required for bacterial differentiation in symbiosis with *Aeschynomene* legumes. *Mol Plant-Microbe Interact* 28: 1155–1166.
 32. Crespo-Rivas, J.C., Guefrachi, I., Mok, K.C., Villaécija-Aguilar, J.A., Acosta-Jurado, S., Pierre, O., et al. (2016) *Sinorhizobium fredii* HH103 bacteroids are not terminally differentiated and show altered O-antigen in nodules of the Inverted Repeat-Lacking Clade legume *Glycyrrhiza uralensis*. *Environ Microbiol* 18: 2392–2404.
 33. Montiel, J., Szűcs, A., Boboescu, I.Z., Gherman, V.D., Kondorosi, E., and Kereszt, A. (2016) Terminal bacteroid differentiation is associated with variable morphological changes in legume species belonging to the Inverted Repeat-Lacking Clade. *Mol Plant Microbe Interact* 29(3): 210–9.
 34. 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: 15244–15249.
 35. Tiricz, H., Szűcs, A., Farkas, A., Pap, B., Lima, R.M., Maróti, G., et al. (2013) Antimicrobial nodule-specific cysteine-rich peptides induce membrane depolarization-associated changes in the transcriptome of *Sinorhizobium meliloti*. *Appl Environ Microbiol* 79: 6737–6746.
 36. Kulkarni, G., Busset, N., Molinaro, A., Gargani, D., Chaintreuil, C., Silipo, A., et al. (2015) Specific hopanoid classes differentially affect free-living and symbiotic states of *Bradyrhizobium diazoefficiens*. *MBio* 6: 1–9.
 37. Oono, R., Schmitt, I., Sprent, J.I., and Denison, R.F. (2010) Multiple evolutionary origins of legume traits leading to extreme rhizobial differentiation. *New Phytol* 187(2): 508–20
 38. Karmakar, K., Kundu, A., Rizvi, A.Z., Dubois, E., Severac, D., Czernic, P., et al. (2019) Transcriptomic analysis with the progress of symbiosis in 'crack-entry' legume *Arachis hypogaea* highlights its contrast with 'infection thread' adapted legumes. *Mol Plant-Microbe Interact* 32(3):271–285.
 39. Kereszt, A., Mergaert, P., Montiel, J., Endre, G., and Kondorosi, E. (2018) Impact of plant peptides on symbiotic nodule development and functioning. *Front Plant Sci.* 9:1026
 40. Trujillo, D.I., Silverstein, K.A.T., and Young, N.D. (2019) Nodule-specific PLAT domain proteins are expanded in the *Medicago* lineage and required for nodulation. *New Phytol* 222: 1538–1550.
 41. Gourion, B., Sulser, S., Frunzke, J., Francez-Charlot, A., Stiefel, P., Pessi, G., et al. (2009) The PhyR- σ EcfG signalling cascade is involved in stress response and symbiotic efficiency in *Bradyrhizobium japonicum*. *Mol Microbiol* 73: 291–305.
 42. Roux, B., Rodde, N., Jardinaud, M.F., Timmers, T., Sauviac, L., Cottret, L., et al. (2014) An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J* 77(6):817–37.
 43. Sen, D. and Weaver, R.W. (1981) A comparison of nitrogen-fixing ability of peanut, cowpea and siratro plants nodulated by different strains of *Rhizobium*. *Plant Soil* 60: 317–319.
 44. Oono, R. and Denison, R.F. (2010) Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol* 154: 1541–1548.

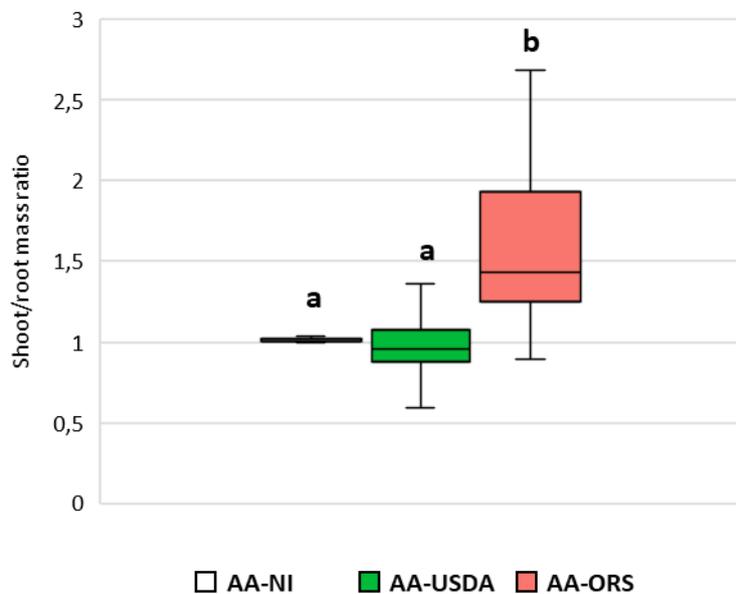
45. Kazmierczak, T., Nagymihaly, M., Lamouche, F., Barriere, Q., Guefrachi, I., Alunni, B., et al. (2017) Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium* strains. *Mol. Plant Microbe Interact.* 30:399–409.
46. Lamouche, F., Bonadé-Bottino, N., Mergaert, P., and Alunni, B. (2019) Symbiotic efficiency of spherical and elongated bacteroids in the *Aeschynomene-Bradyrhizobium* symbiosis. *Front. Plant Sci* 10: 377.
47. Regensburger, B. and Hennecke, H. (1983) RNA polymerase from *Rhizobium japonicum*. *Arch Microbiol* 135: 103–109.
48. Giraud, E., Hannibal, L., Fardoux, J., Verméglio, A., and Dreyfus, B. (2000) Effect of *Bradyrhizobium* photosynthesis on stem nodulation of *Aeschynomene sensitiva*. *Proc Natl Acad Sci U S A* 97: 14795–14800.
49. Chappelle, E., Alunni B, Malfatti P, Solier L, Pedron J, Kraepiel Y, Van Gijsegem F. (2015) A straightforward and reliable method for bacterial in planta transcriptomics: application to the *Dickeya dadantii/Arabidopsis thaliana* pathosystem. *Plant J* 82:352–362.
50. Azani, N., Babineau, M., Bailey, C.D., Banks, H., Barbosa, A.R., Pinto, R.B., et al. (2017) A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny – The Legume Phylogeny Working Group (LPWG). *Taxon* 66: 44–77.
51. Brottier, L., Chaintreuil, C., Simion, P., Scornavacca, C., Rivallan, R., Mournet, P., et al. (2018) A phylogenetic framework of the legume genus *Aeschynomene* for comparative genetic analysis of the Nod-dependent and Nod-independent symbioses. *BMC Plant Biol* 18: 333.
52. Langella O, Valot B, Jacob D, Balliau T, Flores R, Hoogland C, Joets J, Zivy M. (2013) Management and dissemination of MS proteomic data with PROTIcDb: example of a quantitative comparison between methods of protein extraction. *Proteomics* 13(9):1457-66.
53. Kessner D, Chambers M, Burke R, Agus D, Mallick P. (2008) ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* 24(21):2534-6.
54. Craig R, Beavis RC. (2004) TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20(9):1466-7.
55. Langella O, Valot B, Balliau T, Blein-Nicolas M, Bonhomme L, Zivy M. (2017) XITandemPipeline: A tool to manage sequence redundancy for protein inference and phosphosite identification. *J Proteome Res.* 16(2):494-503.
56. Delmotte N, Mondy S, Alunni B, Fardoux J, Chaintreuil C, Vorholt JA, Giraud E, Gourion B. (2014) A proteomic approach of *Bradyrhizobium/Aeschynomene* root and stem symbioses reveals the importance of the *fixA* locus for symbiosis. *Int J Mol Sci.* 15(3):3660-70. doi: 10.3390/ijms15033660.
57. Su, F., Gilard, F., Guérard, F., Citerne, S., Clément, C., Vaillant-Gaveau, N., and Dhondt-Cordelier, S. (2016) Spatio-temporal responses of *Arabidopsis* leaves in photosynthetic performance and metabolite contents to Burkholderia phytofirmans PsJN. *Front Plant Sci* 7: 1–15.
58. Médigue, C., Calteau, A., Cruveiller, S., Gachet, M., Gautreau, G., Josso, A., et al. (2019) MicroScope- an integrated resource for community expertise of gene functions and comparative analysis of microbial genomic and metabolic data. *Brief Bioinform* 20: 1071-1084.
59. Ledermann, R., Bartsch, I., Remus-Emsermann, M.N., Vorholt, J.A., and Fischer, H.M. (2015) Stable fluorescent and enzymatic tagging of *Bradyrhizobium diazoefficiens* to analyze host-plant infection and colonization. *Mol Plant-Microbe Interact* 28: 959–967.

60. Ducret, A., Quardokus, E. M., & Brun, Y. V. (2016). MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nature Microbiology*, 1(7), 16077.
61. Beroual, W., and Biondi, E.G. (2019) A new factor controlling cell envelope integrity in Alphaproteobacteria in the context of cell cycle, stress response and infection. *Mol Microbiol.* 111(3):553-555.

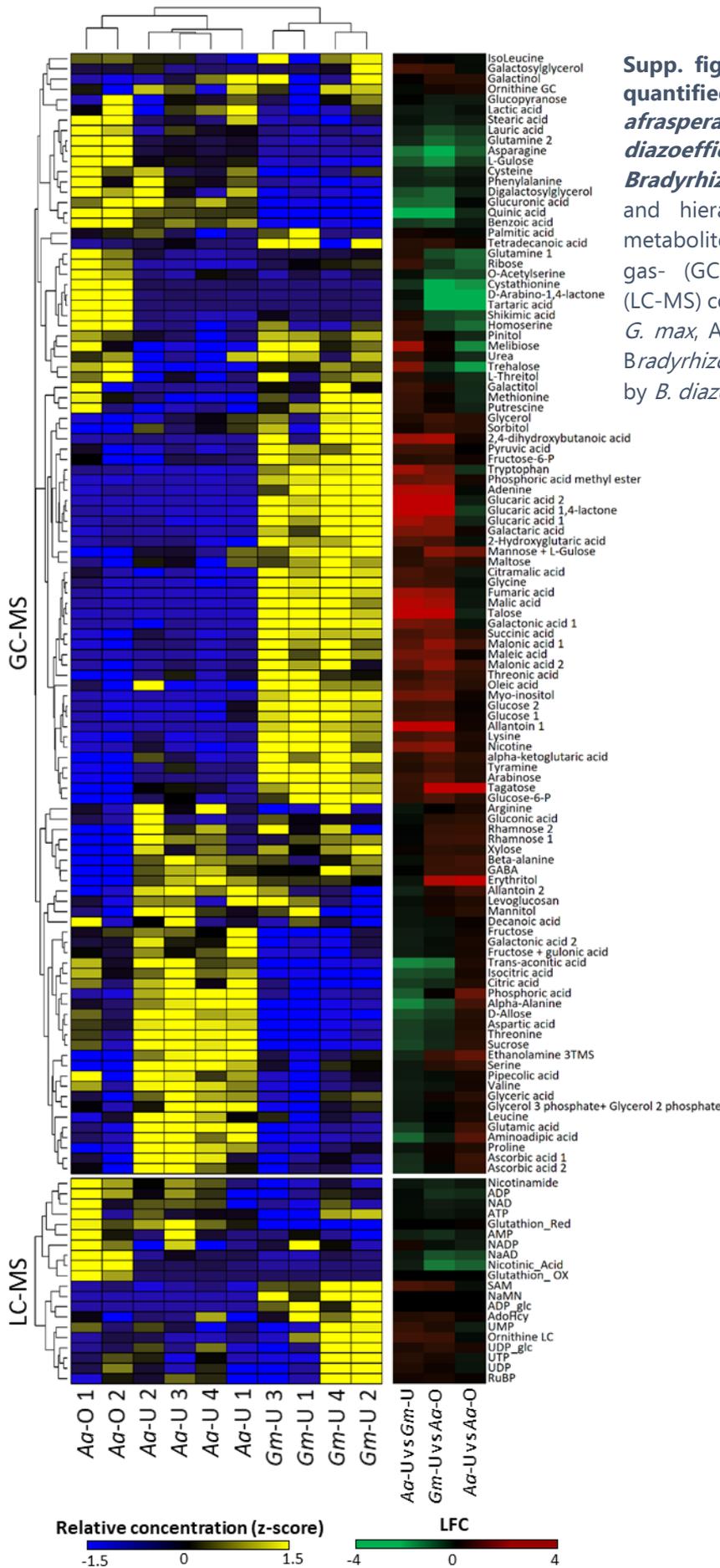
xii. Supplementary figures and tables



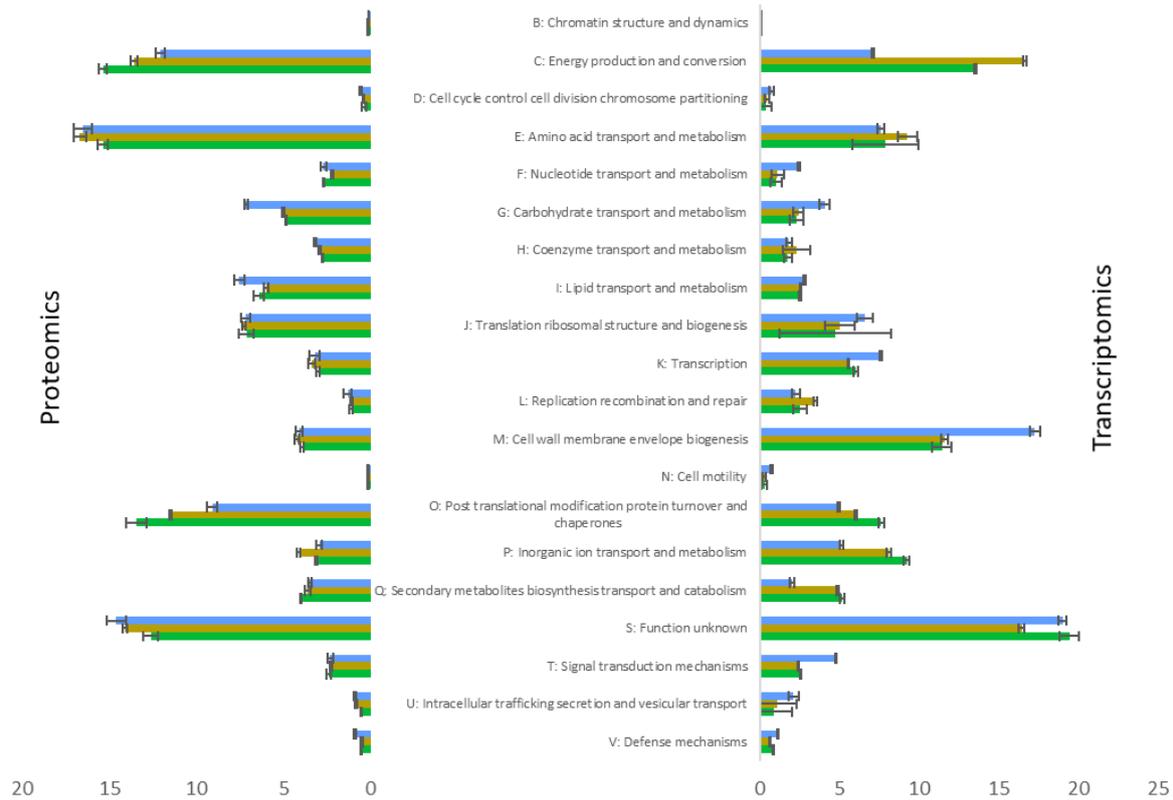
Supp. figure A1 | Nitrogen and carbon content in aerial parts of the plants were determined by elemental analysis. GM: *G. max*, AA: *A. afraspera*, ORS: inoculated by *Bradyrhizobium* sp. ORS285, USDA: inoculated by *B. diazoefficiens* USDA110, NI: Non-inoculated plants.



Supp. figure A2 | Nutritional status of 14 dpi plants determined by the shoot/root mass ratios. AA: *A. afraspera*, ORS: inoculated by *Bradyrhizobium* sp. ORS285, USDA: inoculated by *B. diazoefficiens* USDA110, NI: Non-inoculated plants. Letters represent significant differences after t-test or ANOVA and post hoc Tukey tests ($p < 0.05$).



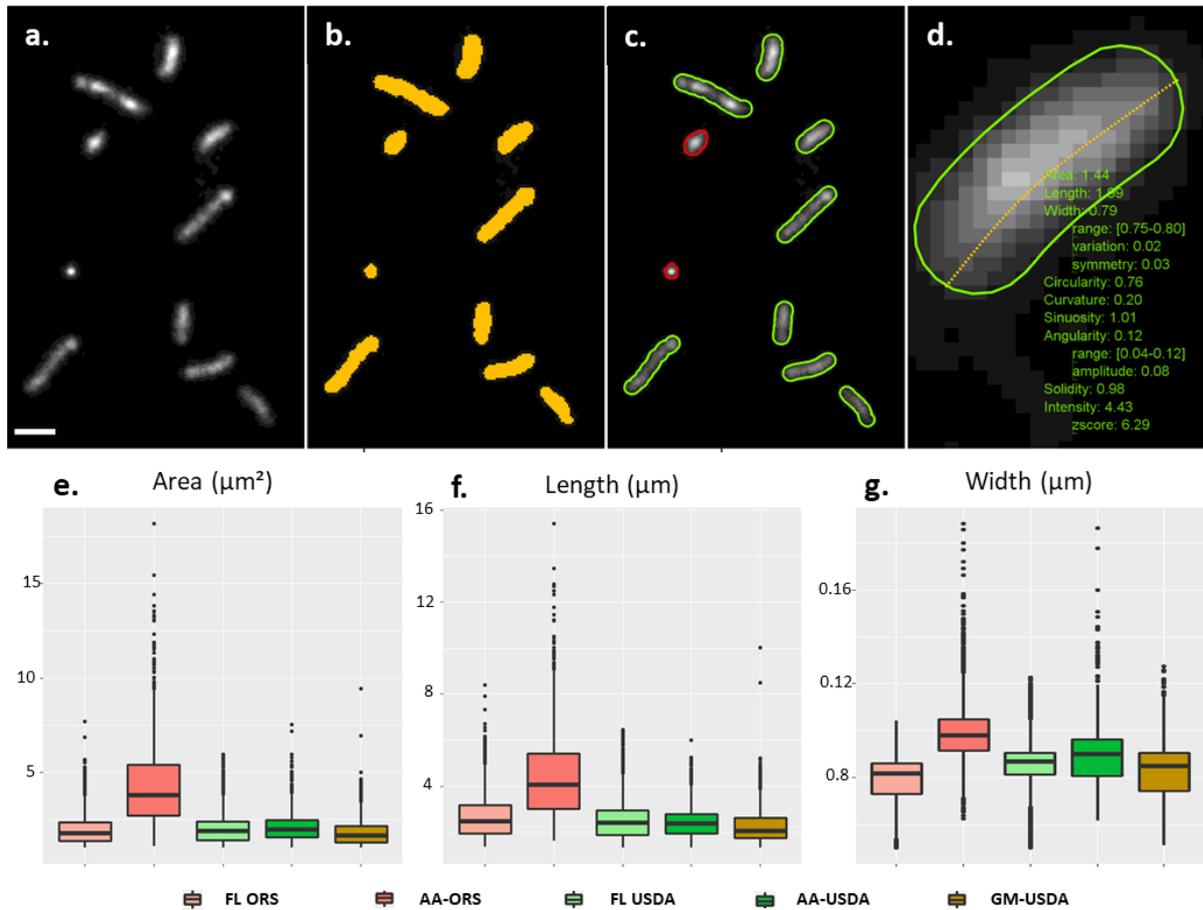
Supp. figure A3 | Overview of the 129 quantified metabolites in *G. max* and *A. afraspera* whole nodules elicited by *B. diazoefficiens* USDA110 or *Bradyrhizobium* sp. ORS285. Heatmap and hierarchical clustering of the 129 metabolites that were quantified either by gas- (GC-MS) or liquid-chromatography (LC-MS) coupled to mass spectrometry. Gm: *G. max*, Aa: *A. afraspera*, O: inoculated by *Bradyrhizobium* sp. ORS285, U: inoculated by *B. diazoefficiens* USDA110.



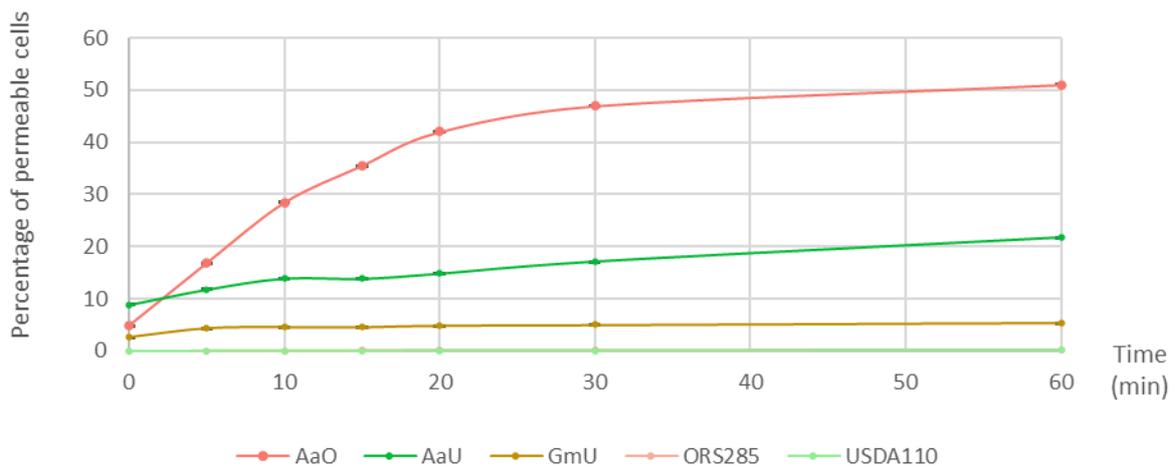
Supp. figure A4 | General overview of the datasets using COG classification. Repartitions of the assigned spectra (left panel) and normalized reads (right panel) among COG classes in the three conditions (blue: bacterial culture, orange: *B. diazoefficiens* USDA110 in *G. max* nodules, green: *B. diazoefficiens* USDA110 in *A. afraspera* nodules).



Supp. figure A5 | Western blot analysis of selected USDA110 proteins in culture and in bacteroids. NifH protein were analyzed by western blots on purified USDA110 bacteroids extracted from soybean and *A. afraspera* nodules 14 dpi. Exponential and stationary phase cultures were used as controls.



Supp. Figure A6 | Analysis of cellular differentiation using automated morphometry. a, b, c & d | Parameters were quantified by image analysis of syto9 stained bacteria and bacteroids using MicrobeJ. The process from raw images (a), segmentation (b), object detection (c) and measurements (d) is depicted with these four panels. **e** | Cell area. **f** | Cell width. **g** | Cell length.



Supp. Figure A7 | Kinetic analysis of bacterial membrane permeability. Kinetics of propidium iodide uptake assays (reflecting membrane permeability) from which data presented in **Figure 5C** were extracted. The PI permeability was measured by flow cytometry over 60 min after treatment on *A. afraspera* nodule extracted USDA110 (AaU) or ORS285 (AaO) bacteroids and *G. max* extracted USDA110 bacteroids at 14 dpi (GmU). Exponential phase bacterial culture of USDA110 and ORS285 were used as controls. Each dot represents three independent measures and error bars represent the standard deviation of the samples.

Table S1 | Genome annotation, transcriptomic and proteomic data of *B. diazoefficiens* USDA110 generated in this study. Description of proteomic and transcriptomic data of USDA110 related conditions. DESeq2 normalized reads, false discovery rate (FDR) values as well as log₂ fold change (LFC) are used to describe transcriptomic data. On the other side, spectral counting (SC) along with related statistical indicators, Tukey statistical test result and p-value depict the proteomic data.

Table S2 | Expression analysis of selected *B. diazoefficiens* USDA110 regulons and stimulons. Detailed analysis of the previously determined regulons and stimulons of USDA110 based on our transcriptomic data. A given regulon/stimulon was considered differentially regulated when $\geq 40\%$ of the corresponding genes were differentially expressed in our conditions.

Table S2 | List of the 3725 orthologous genes shared by *B. diazoefficiens* USDA110 and *Bradyrhizobium* sp. ORS285 with their corresponding expression level in rich medium and in *A. afraspera* nodules. This dataset was obtained after a Phyloprofile analysis on Mage Microscope website and was used to generate the **Figure 4**. Normalized read counts are shown together with the corresponding LFC and FDR as determined by DESeq2.

Note: All tables can be downloaded here:

<https://www.biorxiv.org/content/biorxiv/early/2020/11/28/2020.11.24.397182/DC2/embed/media-2.xlsx?download=true>

xiii. Additional results and discussion

These results are not be part of the final submitted manuscript.

Functional analysis of similarly regulated genes in USDA110 bacteroids from A. afraspera and G. max

The role in bacteroid functioning of many of the discovered bacteroid specific DEG and DAP was not studied before or only in symbiosis with *G. max*. Therefore, loss-of-function mutants were created in a few selected genes to analyze and compare phenotypes in the symbiosis with the two studied host plants. One of them is the gene *fppS_1* (*blr2148*). It belongs to a gene cluster of nine genes (*blr2143-blr2150*), including four cytochrome P450 encoding genes, involved in the biosynthesis of the phytohormone gibberellin (Nett et al., 2017). In this pathway, *fppS_1*, which encodes the geranylgeranyl diphosphate synthase, performs the first step. The cluster is upregulated from background expression levels in culture to very high levels in bacteroids, four proteins, including *fppS_1*, being also over-accumulated *in planta* (**fig. 30a**). The operon is up-regulated similarly as the nitrogenase genes, by the NifA transcriptional regulator and the low oxygen levels in nodules (Hauser et al., 2007). Gibberellin production by bacteroids is intriguing because this hormone has a well-defined role in the establishment of the rhizobium-legume symbiosis (Fonouni-Farde et al., 2016). However, we found that the *fppS_1* mutant has no phenotype on nodule formation and functioning either in *G. max* or in *A. afraspera*.

A locus of seven genes (*blr7916-blr7922*) encoding an amidase enzyme and a putative peptide transporter composed of two transmembrane domain proteins, two ATPases and two solute-binding proteins was strongly upregulated in the two bacteroid types, with three protein being also over-accumulated *in planta* (**Fig. 30a; Table S1**). The two genes *blr7921* and *blr7922*, encoding the solute-binding proteins, as well as those encoding the permease and ATP-binding domain of this complex (*blr7918-blr7920*) were mutagenized by deletion or gene interruption. However, none of the mutants displayed a symbiotic phenotype that deviated significantly from the wild type (**Fig. 33**).

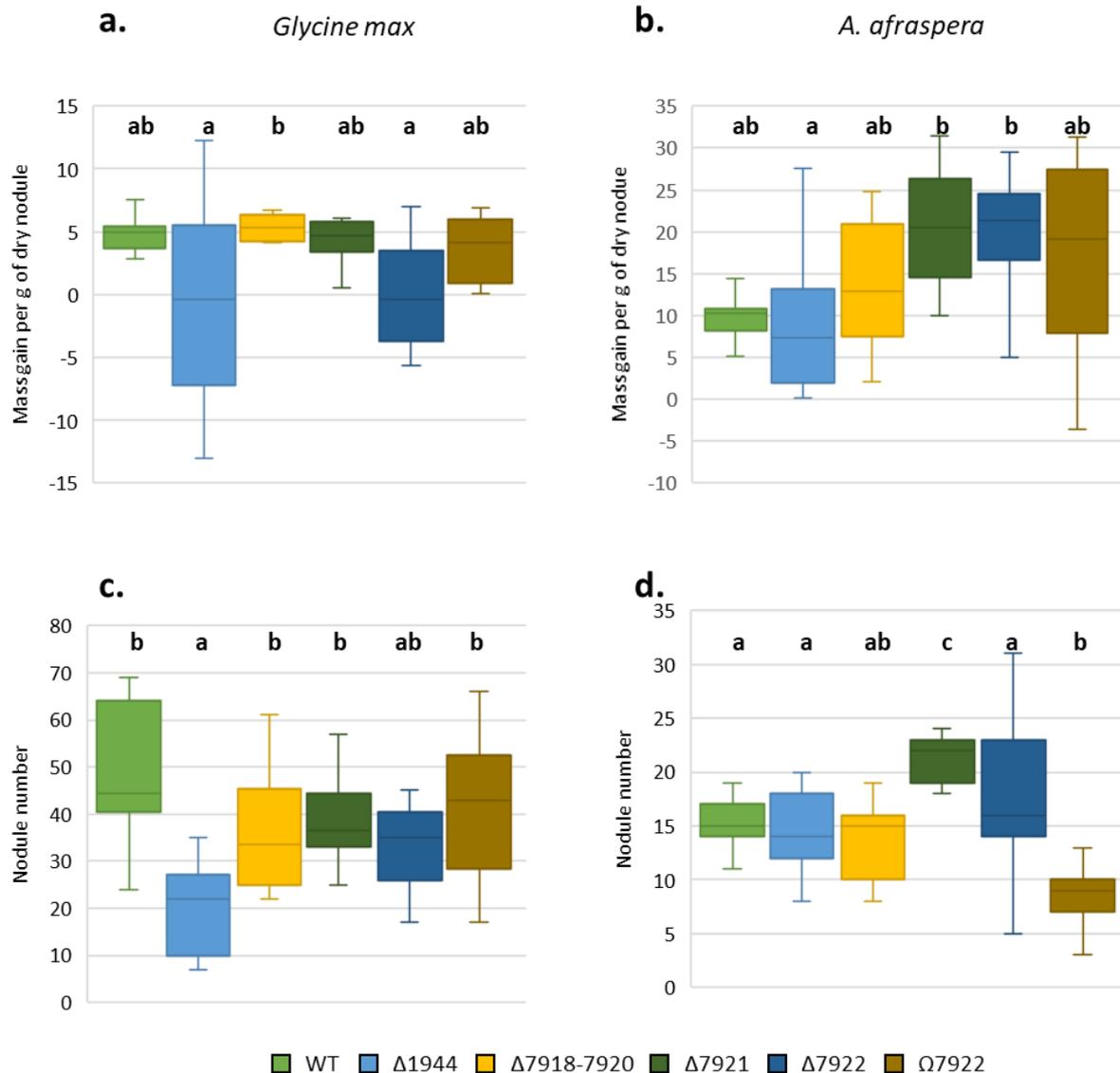


Figure 33 | Functional studies of selected genes displaying no or little symbiotic phenotypes. a, b | Comparative symbiotic mass gain of *B. diazoefficiens* USDA110 mutants compared to the wild type strain in *G. max* (a) and *A. afraspera* (b) nodules. **c, d** | Nodule number elicited by *B. diazoefficiens* USDA110 wild type and mutant strains in *G. max* (c) and *A. afraspera* (d). Letters represent significant differences after ANOVA and post hoc Tukey tests ($p < 0.05$). WT: wild type, Δ : deletion mutant, Ω : pVO155 insertion mutant.

Another mutant in the *bll1944* gene encoding a highly abundant bacteroid-specific outer membrane peptidase was also generated. This gene was not specifically differentially expressed but displayed a higher protein level in soybean than in *A. afraspera* (Table S1). The mutant in *bll1944* had a mildly affected symbiotic phenotype, displaying a lower efficiency and nodule number compared to the wild type in symbiosis with *G. max*, whereas no defects were observed in interaction with the *A. afraspera* host (Fig. 33).

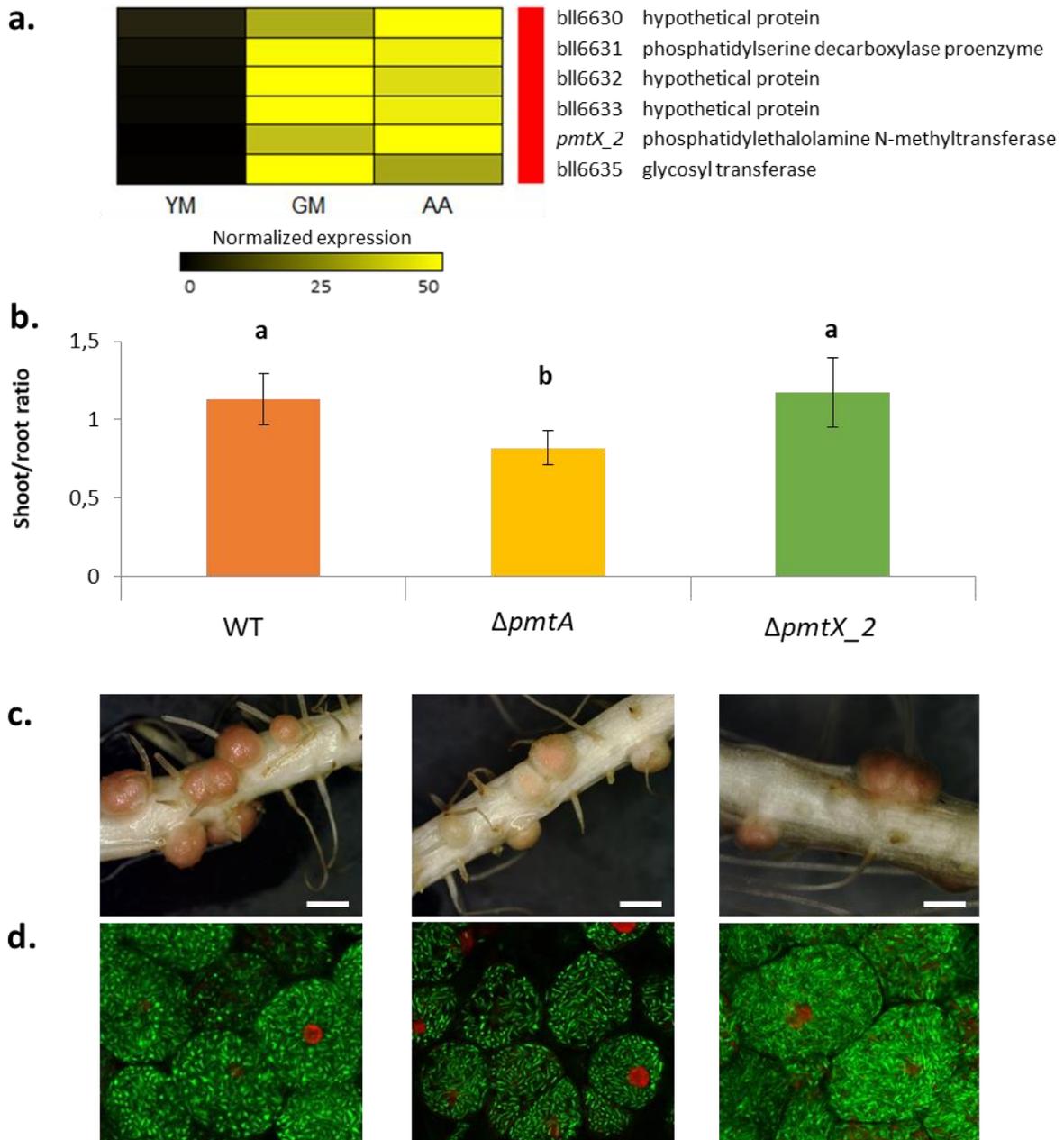


Figure 34 | Functional studies of genes involved in phospholipid synthesis. a | Heatmap of a gene cluster upregulated in both bacteroid conditions (RNA-seq). The right column indicates whether the gene is differentially expressed (red) or not (black) in at least one condition ($fdr < 0.01$ and $|LFC| > 1.58$). **b |** Analysis of the nutritional status of *A. afraspera* plants by shoot/root mass ratio measurements after 14 days after inoculation with $\Delta pmtA$ and $\Delta pmtX_2$ mutants or with the wild type (WT) strain. Letters represent significant differences after ANOVA and post hoc Tukey tests ($p < 0.05$). **c |** Macroscopic images of *A. afraspera* nodules, confirming the symbiotic phenotype of $\Delta pmtA$ mutant previously observed on *G. max*. **d |** Confocal images of live/dead stained symbiotic cells showing a reduced bacteroid density for $\Delta pmtA$ mutant. Scale bars: 0,5 μm (**c**) and 10 μm (**d**).

The transcriptome dataset unveiled a gene cluster involved in phosphatidylcholine synthesis (*blI6630-blI6635*) that is strongly upregulated in bacteroids (**Fig. 34a**). This gene cluster encodes a phosphatidylserine decarboxylase, a phosphatidylethanolamine N-

methyltransferase (PmtX_2) and a glycosyltransferase. The orthologous cluster in *Bradyrhizobium* sp. ORS285 is also upregulated in bacteroid conditions, where cell elongation occurs during TBD (Lamouche et al., 2019a). On the other hand, other genes of the phosphatidylcholine synthesis pathway, *pmtA* (*blr0681*), *pmtX_1* (*bll699A*), *pmtX_3* (*bll8166*), *pmtX_4* (*blr480A*), *pssA* (*blr3797*) and *psd* (*blr3796*) were not differentially regulated in culture and bacteroids. We tested the phenotype of available USDA110 mutants in the *pmtA* (*blr0681*) and *pmtX_2* (*bll6634*) genes, both encoding phosphatidylethanolamine N-methyltransferases (Hacker et al., 2008; Minder et al., 2001). The $\Delta pmtA$ mutant displayed a symbiotic defect in *A. afraspera* nodules, as observed before in *G. max* (Minder et al., 2001). *A. afraspera* plants inoculated with this mutant showed a low shoot/root ratio, suggesting a nutritional stress (**Fig. 34b**). The nodule cells seem to be properly infected with the $\Delta pmtA$ mutant, but the nodules are much less pink than wild type nodules, in agreement with a weakly functional symbiosis (**Fig. 34**). In contrast, no symbiotic defects could be identified on both host plants for the $\Delta pmtX_2$ mutant, despite its high expression in bacteroids (Hacker et al., 2008; **Fig. 34**). In addition, mutants were generated in the bacteroid-enhanced genes *blr1880* (encoding a transcription factor) and *blr1686* (encoding an aminotransferase). However, also for these mutants, no symbiotic defects could be noticed on either plant (data not shown).

Functional analysis of DEG and DAP between A. afraspera and G. max USDA110 bacteroids

A functional analysis of four genes that were differentially regulated between the two hosts, *blr3675* (encoding an alcohol dehydrogenase), *blr7922* (encoding a solute binding protein), revealed once more no striking aberrant symbiotic phenotypes. The genes encoding the Calvin-Benson-Bassham pathway were induced in *A. afraspera* bacteroids, although only to a moderate expression level. The polar mutant $\Omega cbbP$, inactivating *cbbP*, *cbbT*, *cbbA*, *rbcl* and *cbbS*, as well as the mutant $\Delta cbbR$, impaired for the production of the transcriptional regulator CbbR of this gene cluster (Masuda et al., 2017), displayed no aberrant symbiotic phenotypes on both host plants (**Fig. 35**).

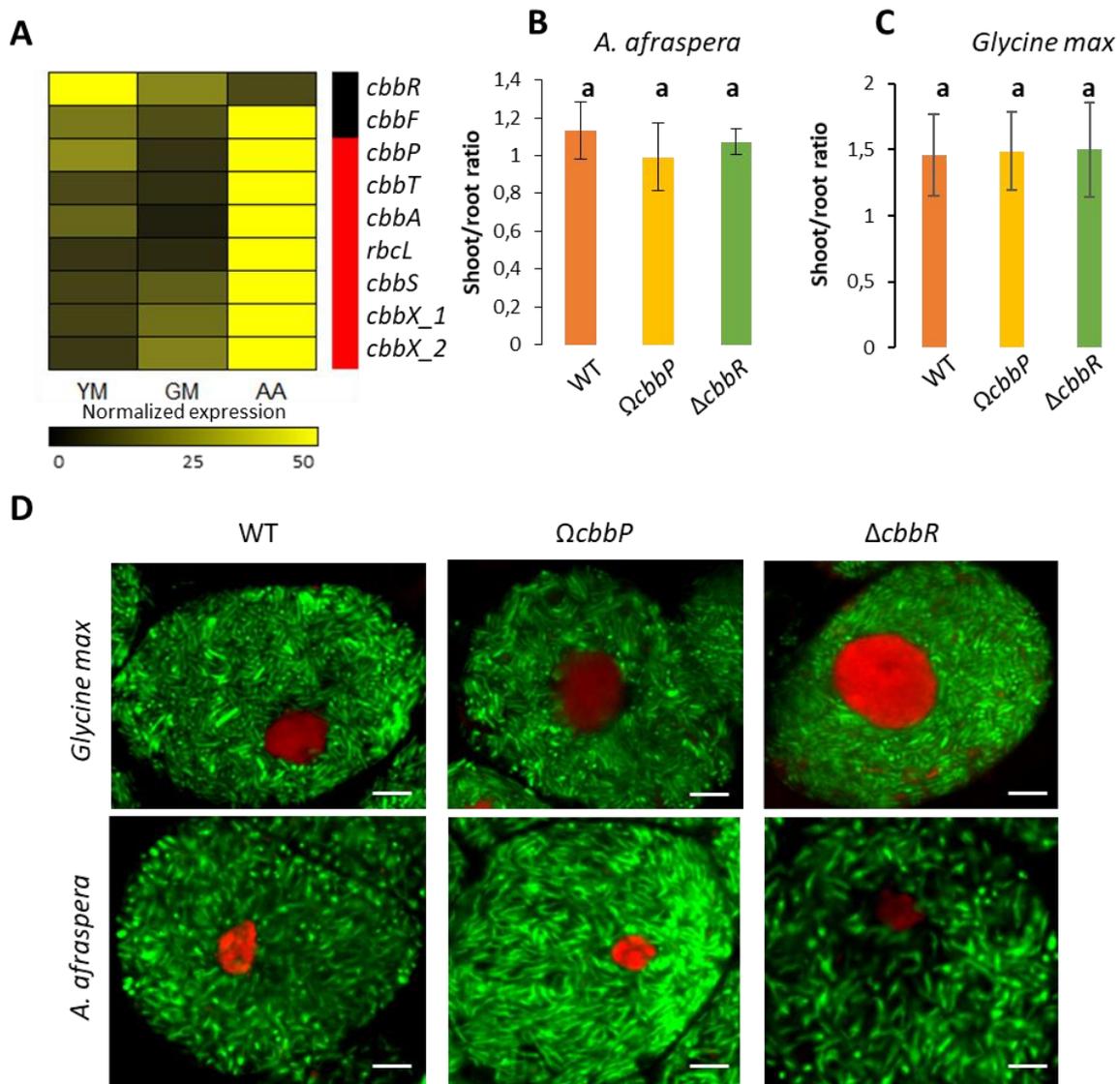


Figure 35 | Functional studies of genes involved in Calvin Benson Bassham pathway. **a** | Heatmap of *cbb* gene expression (RNA-seq), showing their upregulation in *A. afraspera* bacteroids. The right column indicates whether the gene is differentially expressed (red) or not (black) in at least one condition ($fdr < 0.01$ and $|LFC| > 1.58$). **b**, **c** | Analysis of the nutritional status of *A. afraspera* (**b**) and *G. max* (**c**) plants by shoot/root mass ratio measurements 14 days after inoculation with Ω cbbP and Δ cbbR mutants or the wild type (WT) strain. Letters represent significant differences after ANOVA and post hoc Tukey tests ($p < 0.05$). **D**. Confocal images of live/dead stained symbiotic cells displaying an effective symbiosis for the tested mutants. Scale bars: 5 μ m.

Absence of phenotype in target genes mutants - reliability of omics techniques?

As described in the previous sections, various mutants were designed, generated and their phenotype assessed in symbiosis. However, for most of them, no impairment in the symbiotic process was observed and if any the alteration of the interaction efficiency was minor. Different

reasons can explain this phenomenon, based either on genome organization or on target selection criteria.

Redundancy – Besides the methods bacteria have to acquire new functions (described in the introduction), genome reshuffling can lead to the creation of paralogs. These genes, appearing after a genome duplication event and followed by mutation can lead to the acquisition of new functions. However, development of new functions through this mechanism is not systematic, and instead it gives birth to redundant functions. Redundancy helps organisms being resilient in their struggle to survive. Among rhizobia can be found some of the largest bacterial genomes (5 to 10 Mb), sometimes because of the presence of large plasmids, either symbiotic or involved in free-living (Poole et al., 2018). Such large genomes are often correlated with a large panel of genes. For example, USDA110 harbors around 9000 genes on its chromosome. Though this battery of genes is representative of the multiple lifestyles these bacteria have, it might be possible to find redundant functions in such bacteria. A good example of redundancy are pathogenic bacteria. Indeed, these bacteria have recourse to secreted proteins known as effectors to be virulent, usually by shutting down host defenses. It is usual to find molecular, target, or pathway redundancy in such proteins, for the bacteria to maximize their impact on the host. In some cases, redundancy has been shown to be associated with a variation of expression levels to compensate for the loss of one or the other redundant gene (Peng, 2019). Regarding, the functions we targeted, it is possible some of them had some form of redundancy, though having checked for homologous genes in the very same species. To palliate this problem, systematic mutation of redundant genes should be performed to unveil their function. In accordance with the gene expression compensation (Peng, 2019), expression levels of homologs genes should be compared between the WT strain and a mutant in a supposedly redundant function.

Criteria for target choice – Using transcriptomic and/or proteomics data for target selection in sight of functional studies has some drawbacks. The criteria of choice to do so is the difference of expression levels or protein content produced from a given gene. However, differential expression or protein accumulation in set of condition does not necessarily mean they are exclusively so in these conditions. It has been shown a differentially expressed gene is not always essential in a condition such as was described by Giaever et al. in 2002. In this study,

single mutant in almost every genes of *Saccharomyces cerevisiae* were generated. Their fitness in different media was compared to the transcriptomic profiles of the WT strain in same culture conditions. They found that, depending on the culture condition, only 7% or less of upregulated genes were found essential, and many of the genes, which exhibited a fitness decrease, were not differentially regulated. Other more recent studies balanced these results, by comparing transcriptomes with genes essentiality assessed by a Tn-seq (transposon sequencing) experiment (Fang et al., 2013). Tn-seq is a functional genomic technique based on transposon insertion sequencing methods (TIS). It allows functional characterization of the whole set of a bacterium's genes through mass sequencing of the library of transposon insertion mutants (more information in **ANNEX B**). With this study, it was observed that essential genes are overall more induced than non-essential genes. However, in this study was analyzed the base essentiality of the bacterium, or functions which are necessary for a bacterium's life in any situation. This comprises primary metabolism or cell-cycle as an example. In our case, we are looking for conditionally essential genes, specific to the survival in the nodule environment. It is likely that depending on the tested conditions the correlation between transcriptomic data and gene essentiality varies. Proteomics criteria for target selection is less distant from the real function compared to that of transcriptomics. However, we still lack information on post-translational regulation that might drive the target selection wrong. Therefore, in bacteria, the use of Tn-seq techniques should be a good alternative to both transcriptomics and proteomics.

If we add up all these reasons together, we might have a good explanation of why no selected genes had altered phenotype in symbiosis. Moreover, as knowledge expands on a subject, researchers will be looking for fine-tuning of their favorite function. Therefore, these techniques might reach their limits if used alone. A way to palliate this is to multiply different omics techniques, such as we did in this study, although it does not always solve this problem. Moreover, managing the analysis of multiple datasets coming from different techniques and tested conditions can be difficult.

Modeling as a way to overview the functioning of a metabolism and regulation networks

The best way to integrate multiple omics data is modelling of an organism biology. Since the last decades, and with the evolution of computational tools both on the hardware and software side, modelling has become more and more accessible. Different types of model can be built to simulate a biological system, from microbial communities, gene regulation networks (GRN) and metabolic models (MM). As we are here interested in a single bacterium rather than a community, I will focus on the two last types of models. However, in our symbiotic systems, inferring microbial communities is likely to be relevant as symbiotic bacteria all need to maintain and compete in the soil and the root.

Due to the complexity of multicellular organisms such as animals or plants, GRN inference and MM were used in their proper extent when applied to bacteria. Indeed, some GRN inference exist in plant studies such as for modelling the behavior of the circadian clock (Aalto et al., 2020) or simplified development features (Alvarez-Buylla et al., 2007). Though helping outline a particular process, they do not facilitate overviewing the organism's biology and how various functions may interact with each other during the lifetime of the organism. In organisms with a simpler organization, such as bacteria, these models can prove very powerful. Until now, several model bacteria have seen their metabolism and/or GRN inferred. This is the case of *E. coli* for which a combined GRN inference and MM have been developed (Covert, 2009) or *Bacillus subtilis* with independent GRN inference (Faria et al., 2016) and MM (Massaiu et al., 2019). On the rhizobia side, several MM have been proposed such as for *B. diazoefficiens* USDA110 (Y. Yang et al., 2017) and *S. meliloti* 1021 (DiCenzo et al., 2016). The *S. meliloti* model was reused in a more ambitious project where MM of a whole nodule, plant and bacteria included was built (DiCenzo et al., 2019 - BioRxiv) and called ViNE (Virtual Nodule Environment). Therefore, in the future we might try to rely on these tools to help predict functions of interest during symbiosis. In other biological models, MM have been shown to help prediction of essential functions such as in malaria disease caused by *Plasmodium* protozoan (Stanway et al., 2019). Conveniently, besides *in silico* predictions, transcriptomic data can be projected onto MM in order to help overviewing how the gene regulation interacts with

the metabolism. These tools might therefore be of great help for a deeper understanding of our but also future transcriptomic assays on these bacteria.

Thanks to a collaboration with George DiCenzo, who built both *S. meliloti* 1021 and ViNE MMs, as well as Marco Fondi's team (COMBO – Firenze – Italy), we began the building of a MM for *Bradyrhizobium sp.* ORS285 (**ANNEX C**). This model would help comparing USDA110 MM to ORS285 MM in order to better understand the behavior of USDA110 in this ill-adapted symbiosis. Whereas MM would not have helped for transcription factors, among the functions we try to mutate were metabolite processing enzymes (*fpps_1*, phosphatidyl choline synthesis or Calvin-Benson-Bassham genes) for which this model could have helped us. Moreover, in a previous study based on transcriptomic analyses, we also struggled to find essential functions (Lamouche et al., 2019a). This study was essentially based on ORS285 behavior in different host plants and therefore this MM could help give a second life to this data. Additionally, this model can be used to study bacterial features found in ORS285 such as free-living nitrogen fixation or photosynthesis, which are absent in other symbiotic bacteria such as *S. meliloti* 1021 or USDA110.

B. *Sinorhizobium meliloti* functions required for resistance to the antimicrobial NCR peptides and bacteroid differentiation.

Authors. Quentin Nicoud^{1,*}, Quentin Barrière^{1,*}, Nicolas Busset^{1,*}, Mickaël Bourge¹, Romain Lebars¹, Marie Lecroëll¹, Sándor Jeneš², Atilla Kereszt², Claire Boulogne¹, Eva Kondoros², Tatiana Timtchenko¹, Benoît Alunni¹, Peter Mergaert^{1,#}

¹Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France

²Institute of Plant Biology, Biological Research Centre, 6726 Szeged, Hungary

(*: equal contributions)

i. Foreword

In the introduction of this manuscript, I described a set of bacterial functions that could be linked to NCR peptides. However, whether those functions are truly associated with NCR peptide response is not clear yet. Indeed, though molecular analyses such as structure of envelope components are often well described, the characterization of these functions is often associated with superficial *in planta* phenotyping. In this analysis, we selected some of these functions in order to better characterize their role in the TBD. Moreover, the only bacterial function which is clearly linked to NCR activity today is the *bacA* gene. Considering the large battery of NCR peptides host plants can possess, it seems unlikely that only one function is responsible for handling this stress. Therefore, our goal was to bring new insights on the *S. meliloti* 1021 response to NCR peptides. The targeted genes are the glycosyltransferase *lpsB* involved in LPS core synthesis (Campbell et al., 2002), the VLCFA acyl transferase *lpxXL* involved in the synthesis of LPS lipid-A (Haag et al., 2011) and the global stress response *rpoH1* which also has been described to play part in symbiosis (Barnett et al., 2012; Mitsui et al., 2004).

In addition to these genes, we were interested in an ABC transporter encoded by the *yejABEF* operon. These genes were identified in a Tn-seq analysis as being essential to survive NCR247 treatment (Arnold et al., 2017). This ABC-transporter is composed of two permeases encoded by *yejB* and *yejE*, the ATPase domain YejF and the PBP YejA. The YejABEF transporter

has orthologs in numerous bacteria and has been shown to transport AMPs in *E. coli* (Novikova et al., 2007). Its orthologs in animal pathogens are also important for successful infection. Indeed, *Salmonella enterica* and *Brucella melitensis* show reduced virulence, bad persistence in their host as well as increased sensitivity to diverse cationic AMPs (Eswarappa et al., 2008; Wang et al., 2016). As this transporter shows functional similarity with BacA/SbmA transporters we constructed mutants to assess their importance during symbiosis but also to *in vitro* stresses.

All the *S. meliloti* 1021 mutants mentioned here were tested for their sensitivity to NCR peptides, but also to various membrane stresses. The ability of mutant strains to uptake AMPs was also measured. Afterwards, their ability to interact with *M. sativa* and *M. truncatula* plants was assessed. Multiple cell biology techniques were used to measure the ability of bacteroids to survive in the nodule, terminally differentiate and fix nitrogen. This project which started during my PhD was initiated as a follow up of the team's previous *bacA* functional studies. In this project, I took care of most of the symbiotic interaction phenotyping and related experiments. Complementary data which I generated is currently being analyzed. These additional results will allow us to quantify bacteroid morphology parameters, similarly to what was presented in the previous chapter. Once completed this analysis will be added to the manuscript which we will submit to the mBio journal (submission planned for early-mid December

ii. Abstract

Legumes of the *Medicago* genus form symbiosis with the bacterium *Sinorhizobium meliloti* and develop root nodules housing large numbers of the intracellular symbionts. Members of the Nodule-specific Cysteine Rich peptide (NCRs) family induce the endosymbionts into a terminal differentiated state. Individual NCRs are antimicrobial peptides that have the capacity to kill the symbiont, but the nodule cell environment prevents bacteroids death. Moreover, the bacterial broad-specificity peptide uptake transporter BacA and exopolysaccharides contribute to protect the endosymbionts against the toxic activity of NCRs. Here, we show that other *S. meliloti* functions participate in the protection of the endosymbionts. The studied functions include an additional broad specificity peptide importer encoded by the *yejABEF* genes, lipopolysaccharide modifications mediated by *lpsB* and *lpxXL* as well as *rpoH1*, encoding a stress sigma factor. In vitro, mutants in these genes show a strain-specific increased sensitivity profile against a panel of NCRs and form nodules in which bacteroid differentiation is affected. The *lpsB* mutant nodule bacteria do not differentiate, the *lpxXL* and *rpoH1* mutants form some seemingly fully differentiated bacteroids although most of the nodule bacteria are undifferentiated, while the *yejABEF* mutants form hypertrophied but nitrogen fixing bacteroids. The nodule bacteria of all the mutants have a strongly enhanced membrane permeability, which is dependent on the transport of NCRs to the endosymbionts. Our results suggest that *S. meliloti* relies on a suite of functions including peptide transporters, the bacterial envelope structures, and stress response regulators to resist the stressful assault of NCR peptides in the nodule cells.

iii. Importance

The nitrogen fixing symbiosis of legumes with rhizobium bacteria has a predominant ecological role in the nitrogen cycle and has the potential to provide the nitrogen required for plant growth in agriculture. The host plants allow the rhizobia to colonize specific symbiotic organs, the nodules, in large numbers in order to produce sufficient reduced nitrogen for the plant needs. Some legumes, including *Medicago* spp., produce massively antimicrobial peptides to keep this large bacterial population in check. These peptides, known as NCRs, have the potential to kill the rhizobia but in nodules, they only inhibit the division of the bacteria, which

maintain a high nitrogen fixing activity. In this study, we show that the tempering of the antimicrobial activity of the NCR peptides in the *Medicago* symbiont *Sinorhizobium meliloti* is multifactorial and requires the YejABEF peptide transporter, the lipopolysaccharide outer membrane and the stress regulator RpoH1.

iv. Introduction

Antimicrobial peptides (AMPs) are a diverse group of defense peptides, which are essential mediators of innate immunity in eukaryotes. Their function is to attack and kill unwanted and potentially harmful invading microbes. Animals deficient in a specific AMP or in the regulation of AMP expression are more sensitive to bacterial or fungal infections, while the ectopic expression of AMPs in animals and plants can enhance resistance (1–3).

AMPs are also widely used as essential regulators of bacteria in symbiotic associations (2). In symbiosis, host organisms intentionally maintain bacterial partners and the role of “symbiotic” AMPs is therefore not to eradicate the symbiotic microbes but rather to police them. AMPs regulate the species composition of microbiota by exploiting the different levels of resistance of bacterial species. They control the confinement of the microbiota or of specific endosymbionts to their dedicated location, for example the gut lumen, the external epithelial cell layers or a specific symbiotic organ. They are also proposed to optimize the metabolic integration of the endosymbionts with their hosts (4).

An extreme case of deployment of AMPs for controlling endosymbiont populations, involving hundreds of peptides, is described in the rhizobium-legume symbiosis (2, 5–9). Legumes form a symbiosis with phylogenetically diverse nitrogen-fixing soil bacteria, collectively called rhizobia. This nutritional symbiosis provides reduced nitrogen to the plants, enabling them to grow in nitrogen-poor soils that are normally growth limiting for plants. The symbiosis implies the formation of nodules, specific symbiotic organs, on the roots of the plants. These nodules house the nitrogen-fixing rhizobia, which transfer their ammonia production to the plant in return for the exclusive niche that the plant creates for them in the nodules where they multiply massively from a single or very few infecting bacteria to a population of millions.

After endocytic uptake by the symbiotic nodule cells, the multiplied bacteria reside intracellularly in vesicles called symbiosomes, individually or by a few. The nodule cells and symbiosomes establish the optimal conditions for nitrogen fixation and metabolic exchange with the endosymbionts and at the same time keep them in check. The low oxygen levels prevailing in the symbiotic nodule cells transform the rhizobia into a differentiated physiological state, called bacteroid, which is adapted for nitrogen fixation. Moreover, in certain legume clades like the Inverted Repeat Lacking Clade (IRLC) and the Dalbergioids, the physiological transition of the bacteroids is accompanied with a remarkable terminal differentiation process that is manifested in an irreversible loss of the capacity of bacteroids to divide (10, 11). These terminally differentiated bacteroids have a partially permeabilized cell membrane. They are giant bacterial cells, being either long, wide rods, resulting from polar cell elongation, or large spheres, resulting from an isodiametric growth. The cell enlargement is driven by a switch in the bacterial cell cycle, from a regular cycle of replication and division to a cycle of repeated genome replications without divisions, resulting in polyploid bacteroids.

The terminal differentiation is triggered by a family of effectors, produced by the symbiotic nodule cells, and called Nodule-specific Cysteine-Rich (NCR) peptides (11–13). The NCR gene family is notable in several ways. In *Medicago truncatula*, over 600 NCR genes were identified in the genome, and in several other species of the IRLC clade, many hundreds of them are present as well (12, 14). They are often clustered in the genome and undergo a rapid diversification (15–17). Also, in *Aeschynomene* species of the Dalbergioids, several tens of NCR genes are expressed in nodules (11, 18). A transcriptome analysis in *M. truncatula* found that all the 600 NCR genes are only expressed in nodules and specifically in the symbiotic cells; they are moreover very strongly expressed (19). However, they are activated in different waves during the differentiation of the symbiotic cells and bacteroids, including sets of NCR genes activated at the onset and others at the intermediate or final stages of the differentiation. These temporal profiles indicate that the different NCR genes have specific functions during the bacteroid formation process.

The NCR peptides have features shared with AMPs and at least some NCRs, in particular the cationic ones, can in vitro kill or inhibit the growth of not only the rhizobium symbionts but also many other bacteria and even fungi (20). Their major antibacterial activity results from

their capacity to disturb the integrity of the inner and outer membranes of bacteria, leading to a loss of the membrane potential (13). In addition, NCR peptides have intracellular targets, binding to ribosomes and inhibiting translation (21). However, eliminating the endosymbionts is evidently not what happens in the symbiotic nodule cells and the bacteroids remain despite the high charge of NCRs metabolically active for very long. Possibly, the environment of the symbiotic nodule cells and symbiosomes contribute to temper the antimicrobial activity of the peptides. Importantly, also specific functions of the bacteria themselves are NCR-resistance determinants in the bacteroids.

Resistance to AMPs is essential in pathogens to overcome the innate immune response and establish chronic infection of the host. Pathogen mutants with lower or higher resistance to AMPs lose or gain in pathogenicity, respectively (1–3). Gut commensals require AMP resistance functions to mediate resilience in the changing gut environment (22). Similarly, *Sinorhizobium meliloti*, the symbiont of *Medicago* plants, requires AMP resistance proteins for chronic infection of nodule cells. BacA (known as BclA in the bradyrhizobia infecting the NCR-producing *Aeschynomene* (23, 24)) is a peptide transporter required to counter the NCR peptides inside the symbiotic nodule cells (25). *S. meliloti bacA* mutants are hypersensitive to the antimicrobial NCR peptides. They induce nodules and infect their symbiotic cells seemingly normally, but the mutants die as soon as they are released in the symbiotic cells. This death can be avoided by blocking NCR transport to the infecting rhizobia in the *M. truncatula dnf1* mutant (25). BacA and BclA proteins are peptide transporters with a broad spectrum of peptide import activity (23, 24, 26). They can promote the uptake of NCR peptides, suggesting that the transport of NCR peptides through BacA or BclA provides resistance by redirecting them away from the bacterial membrane, thereby limiting membrane damage. Exopolysaccharide (EPS) is another known factor of *S. meliloti* that helps the endosymbionts to withstand the NCRs (14, 27). This negatively charged extracellular polysaccharide traps the cationic AMPs, reducing their effective concentration in the membrane vicinity. Finally, some rhizobia possess an NCR-degrading protease that allows them to escape altogether from the bacteroid differentiation process and that enhances their proliferation within the nodules without fixing nitrogen, resulting in a more parasitic lifestyle of these rhizobia in the nodules (28).

Bacterial resistance to AMPs is usually multifactorial (29), suggesting that besides BacA and EPS, additional functions of *S. meliloti* bacteroids contribute to resist the NCRs in the symbiotic nodule cells. The literature on *S. meliloti* is rich in the description of bacterial genes that are required for symbiosis. However, often the reporting on these mutants lacks precise information on their bacteroid phenotype and/or on their sensitivity to NCRs. Moreover, transcriptome and Tn-seq (transposon sequencing) analyses of NCR-treated cells and NCR-protein interaction studies identified a whole suite of additional candidate NCR-responsive functions in *S. meliloti* (21, 30, 31). Together with BacA and EPS, some of these *S. meliloti* functions may contribute to alleviate the NCR stress on the bacteroids. To test this hypothesis, we have selected in the present study three candidate functions and analyzed the phenotype of the corresponding mutants in NCR resistance and bacteroid formation.

v. Results

Sinorhizobium meliloti mutants with enhanced sensitivity to NCR peptides

The *S. meliloti* functions selected in this study include a broad specificity peptide uptake transporter encoded by the *yeyABEF* genes (*SMc02829-SMc02832*), lipopolysaccharide (LPS) modifications mediated by *lpsB* (*SMc01219*) and *lpxXL* (*SMc04268*) as well as *rpoH1* (*SMc00646*), encoding a stress sigma factor. The YeyABEF ABC transporter, composed of the periplasmic binding protein YejA (*SMc02832*), the ATPase YejF (*SMc02829*) and the permeases YejB (*SMc02831*) and YejE (*SMc02830*), was selected on the basis of the fact that this transporter and its *Escherichia coli* ortholog have an overlapping peptide uptake specificity with BacA (known as SbmA in *E. coli* (32)). Moreover, mutants in the orthologous genes of the animal pathogens *Salmonella enterica* serovar *Typhimurium* and *Brucella melitensis* are more sensitive to diverse antimicrobial peptides, have a lower virulence and a reduced capacity to persist in their animal host (33, 34). Finally, and most importantly, a genetic screen by Tn-seq in *S. meliloti* has revealed that the transporter mutant has an increased sensitivity to the peptide NCR247 (30). AMPs bind anionic LPS in the initial stage of interaction with the bacterial membrane and subsequently insert themselves in the lipid layer composed of the lipid A moiety of LPS. Consequently, LPS structure is one of the major determinants of AMP resistance and sensitivity in Gram-negative bacteria (29, 35). The selected genes *lpsB* and *lpxXL* encode

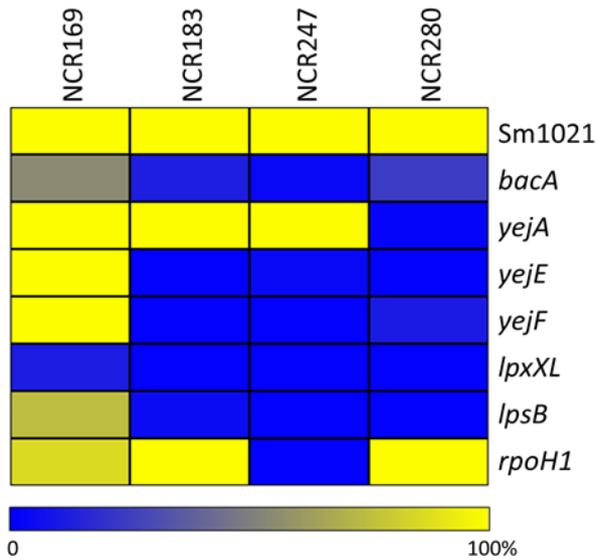


Figure 36 | Sensitivity profile of *Sinorhizobium meliloti* strains to a panel of NCR peptides. The heatmap shows the relative survival, expressed in %, of the mutant strains compared to the wild type survival set at 100% for each peptide treatment.

respectively a glycosyltransferase involved in the synthesis of the LPS core and a very-long-chain fatty acid acyl transferase involved in the biosynthesis of lipid A. Mutants in these genes are affected in resistance to AMPs and in symbiosis (36, 37). Finally, the *rpoH1* gene is a global stress regulator in *S. meliloti* (38). The gene as well as its target genes are upregulated in NCR247 treated cells (31, 39). An *rpoH1* mutant is also affected in symbiosis (40). These selected genes are expressed in nodules, with peak expression in different regions of the nodule where bacteria infect plant cells, undergo the differentiation process, or fix nitrogen (**Supp. Fig. B1**).

Mutants in these candidate genes were constructed (*yejA*, *yejE*, *yejF*, *rpoH1*) or were obtained from previous studies (*lpsB*, *lpxXL*). Log-phase bacteria were treated with a small panel of NCR peptides that were previously shown to have antimicrobial activity (14). These tested peptides were NCR169, NCR247, NCR247 and NCR280, which displayed three different expression patterns in the nodule tissues (19). The NCR280 gene is expressed in the younger nodule cells, the NCR169 gene in the older cells while the two other peptide genes have an intermediate expression pattern (**Supp. Fig. B1**). The bacterial survival after treatment was determined by determining the colony-forming units. The selected mutants were tested along with the wild-type strain and the *bacA* mutant, which was previously shown to be hypersensitive to NCRs (25). The four tested NCR peptides had a strong antimicrobial activity against the wild-type strain, which displayed a survival rate ranging from 8% to 0.03% depending on the tested peptide (**Table S1**). In agreement with previous results, the *bacA*

mutant was hypersensitive to the four peptides (**Table S1; Fig. 36**). Interestingly, the newly analyzed mutants all displayed a higher sensitivity to at least one of the peptides compared to the wild type (**Table S1; Fig. 36**). The *lpxXL* mutant was more sensitive to the four peptides, the *lpsB*, *yejE* and *yejF* mutants were more sensitive to NCR183, NCR247 and NCR280. The *yejA* mutant was more sensitive to NCR280 and the *rpoH1* mutant was more sensitive to NCR247. The differential response towards peptide NCR247 of the *yejA* mutant on the one hand and the *yejE* and *yejF* mutants on the other hand, corresponds to the previously described Tn-seq screen with this peptide (30). Taken together, this analysis indicated that each mutant displayed a specific sensitivity profile to the panel of tested peptides.

Nodule formation by NCR-sensitive Sinorhizobium meliloti mutants

Next, we tested whether the selected *S. meliloti* genes are required for the establishment of a functional symbiosis with *M. truncatula*. The symbiotic phenotype of these mutants was compared with the wild-type strain and the *bacA* mutant. Macroscopic inspection of the root system of plants inoculated with the wild type and seven mutants (**Supp. Fig. B2**) revealed that mutants in the *yejA*, *yejE*, *yejF* and *lpxXL* genes formed nodules with a similar aspect as the functional nodules formed by the wild-type strain, recognizable by their elongated shape and pink color resulting from abundant leghemoglobin production. On the other hand, the *bacA*, *lpsB* and *rpoH* mutants formed abnormal looking nodules that were small and white, symptomatic for the absence of nitrogen fixation (Fix-) and in agreement with previous descriptions (25, 36, 40).

The nitrogen fixation activity of the nodules was measured directly with the Acetylene Reduction Assay (ARA) (**Supp. Fig. B3a**). As expected from the macroscopic observations, nodules infected with the *bacA*, *lpsB* and *rpoH1* mutants were non-functional (Fix-). The other mutants, in the *yejA*, *yejE*, *yejF* and *lpxXL* genes, displayed nitrogen fixation activity (Fix+), again in agreement with the visual features of the nodules. Nevertheless, the *yejE*, *yejF* and *lpxXL* mutants had a reduced activity compared to the wild type.

The histological organization of the nodules formed by the mutants, and in particular the formation of infected symbiotic cells and the viability of the bacteria they contain, was analyzed using fluorescence staining of nodule sections with a mixture of the dyes propidium

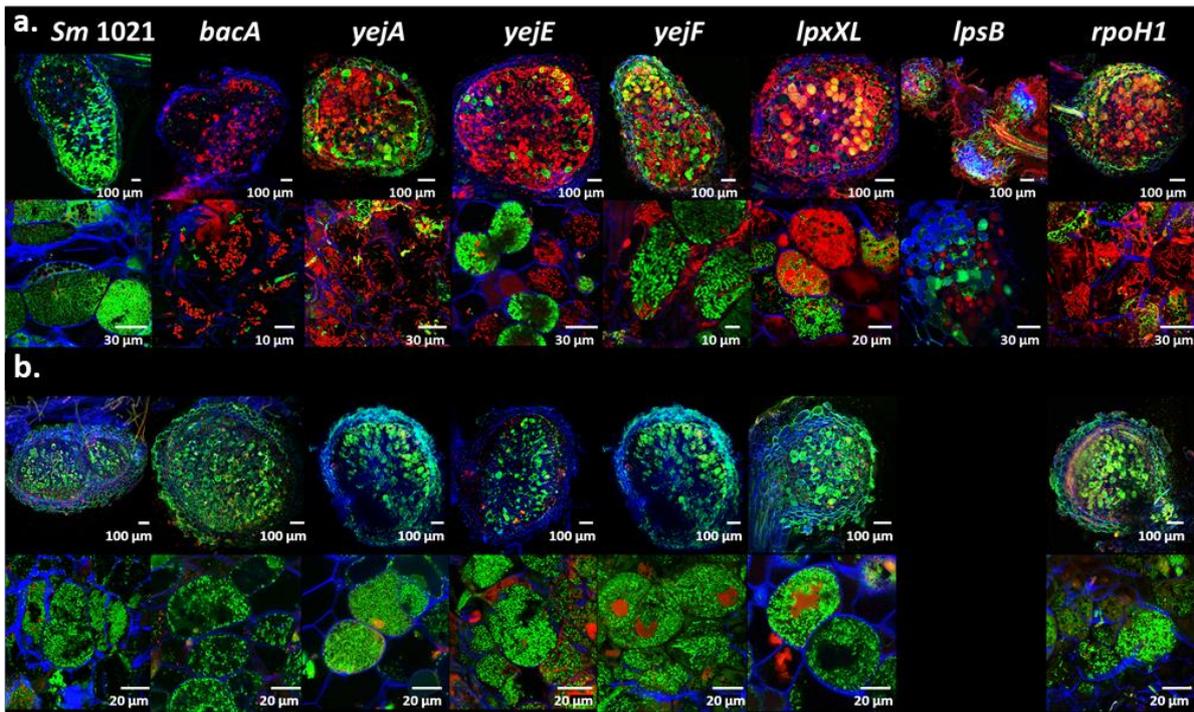


Figure 37 | Symbiotic phenotype of *Sinorhizobium meliloti* 1021 mutants during symbiosis with *Medicago truncatula* wild type and *dnf1* mutant. a | Bacteroid viability determined by live-dead staining of nodule sections and confocal microscopy in wild-type *M. truncatula*. Top row images, full nodule sections; Bottom row images, enlarged images of symbiotic cells. Scale bars are indicated in each panel. b | Membrane permeability of *S. meliloti* 1021 mutants in nodules of the *M. truncatula dnf1* mutant. Nodule phenotype at 21 days post inoculation. The membrane characteristics of the nodule bacteria was determined by live-dead staining of nodule sections and confocal microscopy in the *M. truncatula dnf1* mutant. Top row images, full nodule sections; Bottom row images, enlarged images of symbiotic cells. Scale bars are indicated in each panel.

iodide, syto9 and calcofluor-white and confocal microscopy observation (**Fig. 37a**). This staining procedure highlights the bacteroids and nodule bacteria with a green fluorescence signal (syto 9) when their membranes are well preserved and with a red fluorescence (propidium iodide) when their membranes are highly permeable. As previously reported, wild-type nodules formed symbiotic cells infected with green-labelled elongated bacteroids, while the nodules infected with the *bacA* mutant contained symbiotic cells carrying small undifferentiated bacteria staining red (25).

Contrary to what we expected from the macroscopic inspection of the nodules and their ability to fix nitrogen, the *yejA*, *yejE* and *yejF* mutant bacteroids morphology were substantially altered compared to wild type bacteroids. A high proportion of them were red stained indicating that their membranes were strongly permeabilized. Nevertheless, other host cells contained bacteroids stained in green by the syto9 dye. LpxXL is known to be important, but

not essential to *S. meliloti* during symbiosis with alfalfa (*Medicago sativa*) (41). The mutant forms hypertrophied bacteroids (37). In our experiments, the *lpxXL* mutant displayed elongated bacteroids but mostly permeable to propidium iodide (stained in red). This observation confirmed that *S. meliloti* *LpxXL* is also essential for bacteroid survival to NCR in *M. truncatula* but that it is not crucial for infection and nitrogen fixation. Inside the small bumps elicited by the *lpsB* mutant, no cell seems to be colonized by bacteria as revealed by confocal microscopy. Thus, the *S. meliloti* *lpsB* mutant failed to colonize the host cells and is defective at an earlier stage in infection than the *bacA* mutant, which does form infected plant cells. It was previously reported that the *lpsB* mutant colonizes nodules cells in *M. sativa* suggesting a more severe defect in *M. truncatula* (42, 43).

To test whether the phenotype of the selected mutants was indeed dependent on the host plant, we also analyzed nodulation on *M. sativa* (**Supp. Fig. B3b, Supp. Fig. B4**). A few notable differences were observed. The *yejA* mutant formed on *M. sativa* nodules and bacteroids that had a very similar appearance as the wild type, although nitrogen fixation activity was lower. The *lpsB* mutant formed nodules that developed further than on *M. truncatula*, containing infected nodule cells in which the bacteria were not differentiated and stained heavily with propidium iodide. Finally, the *rpoH1* mutant formed elongated nodules that were pinkish at the tip and showed a slight nitrogen fixation activity. However, these nodules were green at the base, indicating accelerated nodule senescence. Accordingly, the bacteroids in these nodules were elongated stained with propidium iodide, in agreement with the Fix⁻ phenotype of the nodules and the described phenotype of the mutant in alfalfa (40). The other strains had phenotypes that were similar to the above described phenotypes in *M. truncatula*.

Can the increased membrane permeability of the NCR-sensitive Sinorhizobium meliloti mutants be attributed to the NCRs in nodules?

The high membrane permeability of the nodule bacteria of the *bacA* mutant is the result of the action of the NCR peptides in the nodule cells on this NCR-hypersensitive mutant (25). Possibly, the same is true for the other mutants. To test this hypothesis, we made use of the *M. truncatula* *dnf1* mutant, which is defective in a nodule-specific subunit of the signal peptidase complex (44). This mutant cannot process the signal peptides of secretory proteins in the symbiotic

nodule cells and can therefore not transport NCR peptides to the symbiosomes (13). As a result, the wild-type *S. meliloti* nodule bacteria in this mutant are not differentiated. The *bacA* mutant, which is strongly permeabilized and stained by propidium iodide in nodules of wild-type *M. truncatula* plants, is not so in the *dnf1* mutant nodule cells because the bacteria are not challenged anymore with the NCRs (**Fig. 37**) (25). Similarly as the *bacA* mutant, the *yejA*, *yejE*, *yejF*, *lpxXL* and *rpoH1* mutants did not display membrane permeability, as revealed by the absence of propidium iodide staining, in the infected nodule cells of the *dnf1* mutant (**Fig. 37**). This suggests thus that these mutants become membrane-permeabilized by the action of the NCRs and that their symbiotic defects are at least in part due to their hypersensitivity to the NCRs (**Fig. 36**). The *lpsB* mutant did not form detectable nodule-like structures on the *dnf1* roots. We can therefore not conclude on the involvement of the NCR peptides in the symbiotic phenotype of this mutant.

Bacteroid differentiation in the NCR-sensitive Sinorhizobium meliloti

The differentiated bacteroids in *Medicago* nodules are characterized by a very strong cell enlargement and an increase in the ploidy level of the bacteria. These parameters are readily measured by DAPI staining of nodule bacteria and flow cytometry measurement of the fluorescence, reflecting the DNA content, and of the forward and side scatter parameters, reflecting the cell size and the internal cellular complexity (10). The measurement of bacteria extracted from wild-type *M. truncatula* nodules revealed a population of bacteroid cells with a high DNA content, over twenty fold higher than the DNA content in cultured Sm1021 (**Fig. 38**) and increased scatter parameters (**Supp. Fig. B5**). Besides the high ploidy cells, the wild-type nodule samples also contained cells that had flow cytometry parameters similar to cultured bacteria, which could correspond to undifferentiated bacteria present in the nodules, for example bacteria in infection threads or bacteria freshly released in the nodule cells. As previously reported, the nodules infected with the *bacA* mutant did not contain differentiated bacteria (**Fig. 38, Supp. Fig. B5**) (24). The nodule bacteria of the *lpsB* mutant had a profile very similar to the *bacA* mutant (**Fig. 38, Supp. Fig. B5**), confirming the complete absence of differentiated bacteria in the nodules of this mutant. Nodules infected with the *rpoH1* mutant had mostly undifferentiated bacteria although a small amount of fully differentiated cells was

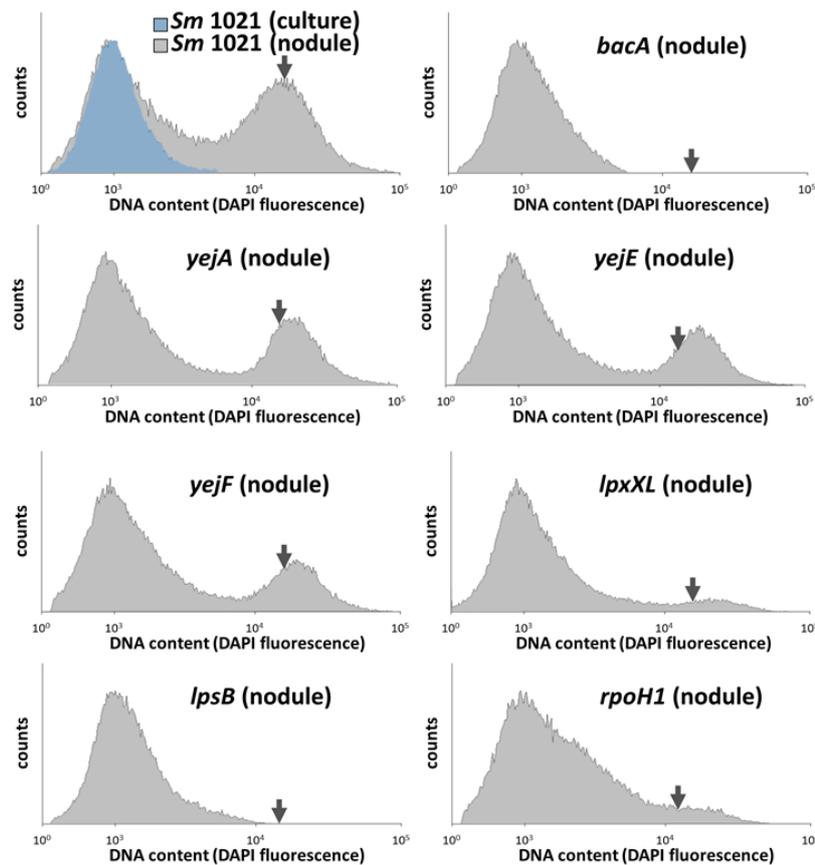


Figure 38 | DNA content in nodule bacteria in *Medicago truncatula*. Flow cytometry analysis of the DNA content of bacteria in culture or isolated from nodules infected with the indicated strains and stained with 4',6-diamidino-2-phenylindole (DAPI). The cell counts (y-axes) are represented in function of the DAPI fluorescence (x-axes). The arrow in each graph indicates the mean DNA content of wild type bacteroids as in the upper left panel.

detected, as well as cells in an intermediate stage. Also, the nodules of the *lpxXL* mutant contained mostly undifferentiated bacteria with a small fraction of fully differentiated ones present (**Fig. 38, Supp. Fig. B5**). In contrast, the *yejA*, *yejE* and *yejF* mutant nodules contained large numbers of fully differentiated bacteria (**Fig. 38, Supp. Fig. B5**).

It was reported previously that the *lpxXL* mutant formed nodules containing misshapen hypertrophied and larger bacteria than the wild type bacteroids (37). In our study, this difference was detectable in the flow cytometry measurements showing a higher DAPI fluorescence and scatter for the small portion of differentiated bacteria in these nodules (**Fig. 38, Supp. Fig. B5**). Moreover, we noticed that the bacteroids in the nodules infected with the *yejA*, *yejE* and *yejF* mutants displayed similar higher DNA fluorescence and scatter (**Fig. 38, Supp. Fig. B5**), suggesting that also these bacteroids have abnormal morphologies.

These patterns, described here for *M. truncatula* nodules were overall similar in the *M. sativa* host plant although the *rpoH1* mutant showed a higher number of intermediate and fully differentiated bacteroids and the *yejA* mutant had a profile similar to the wild type in nodules at 21 dpi while larger than normal bacteroids were only detected at 32 dpi (**Supp. Fig. B6, Supp. Fig. B7**). These differences corresponded well with the macroscopic and microscopic differences in the nodules between the two host plants (**Fig. 37, Supp. Fig. B2, Supp. Fig. B4**).

Defective bacteroid differentiation in yejE and yejF mutants

To confirm the altered bacteroid morphologies in the *yej* mutants, suggested by the cytometry analysis, the *yejE* and *yejF* mutants were observed at high magnification by confocal, fluorescence and transmission electron microscopy.

Confocal microscopy of *M. sativa* nodule sections showed that the nodule cells infected with the *yejE* or *yejF* mutant contained a very heterogeneous population of abnormal bacteroid morphs, including elongated, spherical, pear-shaped or irregular blob-like cells (**Fig. 39a & b**). These cells contrasted strongly with the narrow, elongated wild type bacteroids. The difference was also strikingly discernable by fluorescence microscopy of preparations of purified nodule bacteria (**Fig. 39b**).

To obtain further details of the structure of these aberrant bacteroids, *yejF* nodule sections were observed by transmission electron microscopy (**Supp. Fig. B8**). Several differences were noticed between the wild type and *yejF* mutant. First, the *yejF*-infected cells accumulated starch granules at their periphery, which were never observed in the wild-type infected cells. Second, the aberrant cell morphologies of the *yejF* nodule bacteria, including the spherical, blob-shaped, and pear-shaped morphs, were confirmed. Third, the cytoplasm and inner membrane of many of the *yejF* nodule cells was retracted, leaving very large intermembrane spaces, which in some cases even developed into vacuoles, entirely surrounded by cytoplasm. Some cells had multiple vacuoles, small and large (**Supp. Fig. B8**). In contrast to the bacteroids, in the cultured bacteria, no differences were observed in the ultrastructure of the wild-type and *yejF* mutant bacteria (**Supp. Fig. B9**). Moreover, the mutants behaved like wild type in unstressed growth in culture or in the presence of hydrogen peroxide, sodium chloride and acid stress (**Supp. Fig. B9**). On the other hand, the *yej* mutants were more

sensitive than wild type to the exposure with the detergent sodium dodecyl sulfate (**Supp. Fig. B9**), indicating that these mutants have a reduced ability to cope with membrane stress. Together, these phenotypes are in agreement with the bacteroid defects resulting from a nodule-specific stress such as one derived from the NCR peptides.

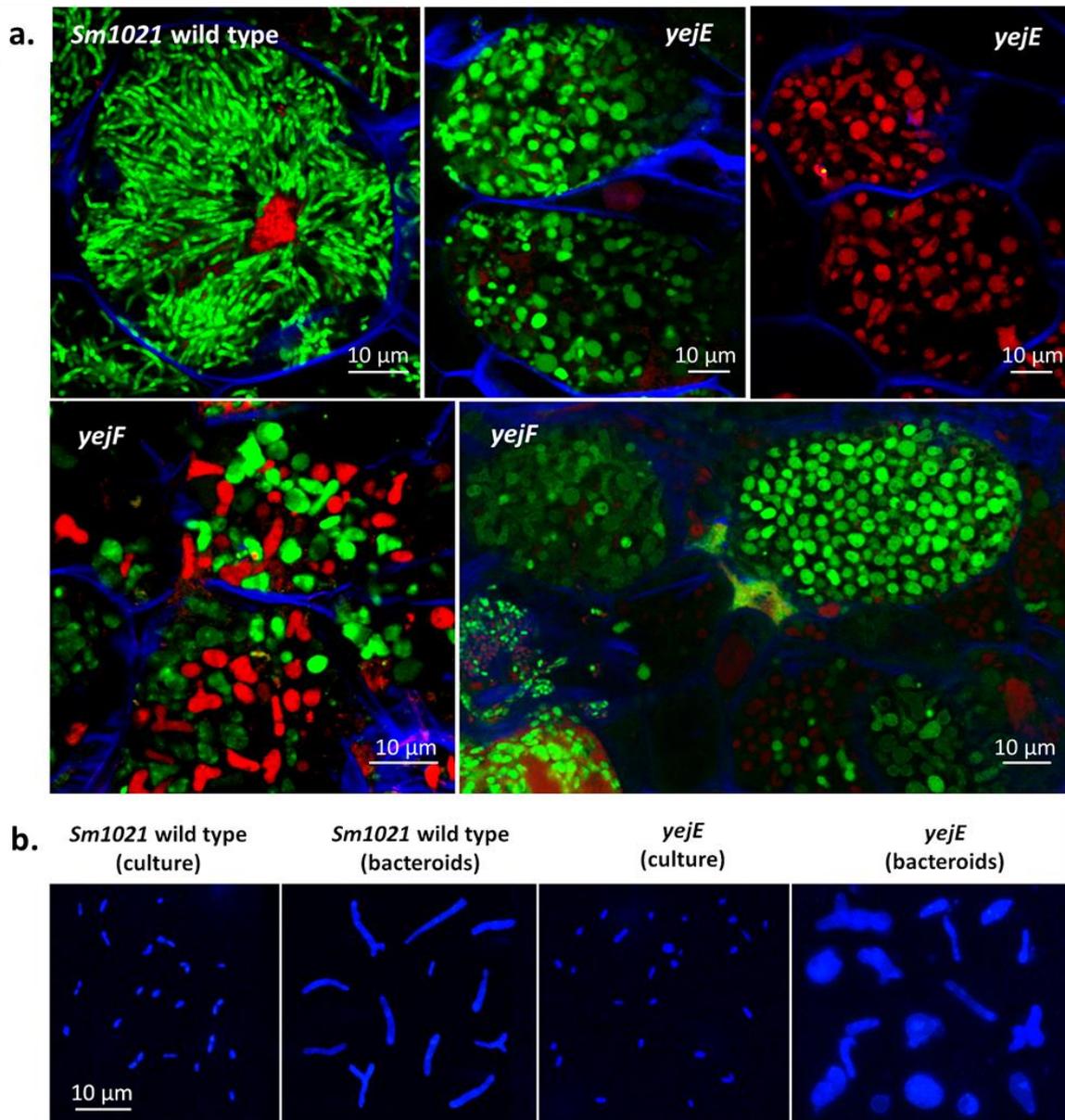


Figure 39 | Bacteroid morphology of *yejE* and *yejF* mutants in *Medicago sativa* nodules. a | Preparations of cultured bacteria or purified *M. sativa* nodule bacteria of the wild type and the *yejF* mutant were observed by fluorescence microscopy. The panels are composite images and the shown individual cells were cut by Photoshop from original images and recombined in a single panel. Each panel is at the same magnification and the scale bar (10 μ m) is indicated in the left panel. b | Sections of nodules infected with the wild type and the *yejE* or *yejF* mutants were stained with a mixture of the dyes propidium iodide, syto 9 and calcofluor-white and observed by confocal microscopy. Scale bars (10 μ m) are indicated in each panel.

Are the hypertrophied bacteroids functional?

The strongly abnormal bacteroids in the *yej* mutants is a priori contradictory with the nitrogen fixation activity of these nodules. To analyze specifically the activity of these aberrant nodule bacteria, we analyzed the *nifH* promoter activity in these bacteroids. The *nifH* gene encodes the nitrogenase reductase subunit of the nitrogenase enzyme and its expression is a marker highlighting nitrogen-fixing bacteroids (45). A GFP gene under the control of the *nifH* promoter located on the plasmid *ppNifH-GFP* was introduced in the wild-type strain and the *bacA* (as a negative control), *yejE* and *yejF* mutants. GFP activity in nodule bacteria of *M. truncatula* (**Fig. 40**) and *M. sativa* (**Supp. Fig. B10**) was analyzed by confocal microscopy and flow cytometry.

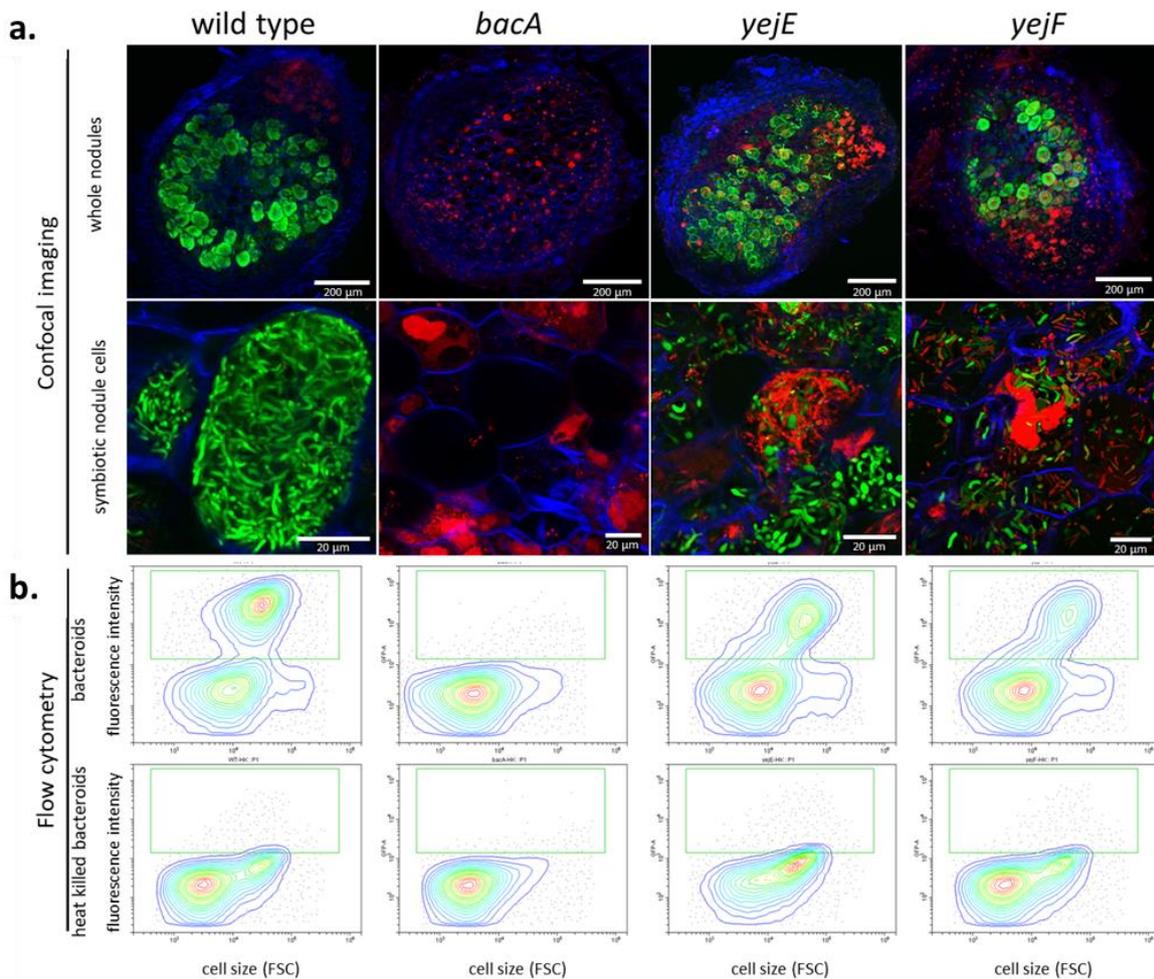


Figure 40 | Nitrogenase expression in the *yejE* and *yejF* mutant bacteroids in *Medicago truncatula* nodules. **a | Confocal microscopy of sections of nodules infected with *Sm1021.ppNifH-GFP*, *bacA.ppNifH-GFP*, *yejE.ppNifH-GFP* or *yejF.ppNifH-GFP* and stained with propidium iodide. **b** | Flow cytometry of GFP levels in differentiated nodule bacteria (upper panels) and heat killed nodule bacteria (lower panels). The square shows the position of the GFP-positive bacteroids.**

Confocal microscopy of nodule sections stained by propidium iodide and calcofluor-white showed that in nodules infected with wild-type Sm1021.*ppnifH-GFP*, all symbiotic cells were filled with *nifH*-expressing cells (**Fig. 40, Supp. Fig. B10**). In contrast, the nodules infected with *yejE.ppnifH-GFP* or *yejF.ppnifH-GFP* contained symbiotic cells with only *nifH*-expressing bacteria, only *nifH*-negative and propidium iodide positive bacteria or a mixture of both. Nevertheless, the microscopy clearly showed that the misshaped bacteroids formed by these mutants can strongly express *nifH* and thus probably can at least temporarily fix nitrogen. In the *bacA* mutant, no GFP signal was detected as expected.

We further analyzed the nitrogenase expression in the *yejE* and *yejF* bacteroids by flow cytometry of the nodule bacteria (**Fig. 40, Supp. Fig. B10**). In wild type, the majority of the differentiated nodule bacteria were strongly GFP positive, confirming that the misshaped bacteroids are functional, even if a fraction of apparently inactive bacteroids were present as well.

Peptide uptake by the YejABEF transporter

In *E. coli*, both YejABEF and SbmA (BacA) mediate the transport of microcin C peptides. Therefore, we tested whether the overlap in substrates of YejABEF and BacA can be extended to other known SbmA/BacA substrates. Bleomycin is a non-ribosomal peptide with DNA damaging activity and thus needs to be internalized, which is mediated by SbmA in *E. coli* and BacA in *S. meliloti* (46). An indirect method to estimate bleomycin uptake by a bacterium and its derivatives is determining their sensitivity to the antibiotic whereby strains that have a reduced uptake activity are more resistant than strains that efficiently uptake the peptide. We found that at 0.5 µg/mL the growth of wild-type *S. meliloti* is nearly completely inhibited while the *bacA* mutant is more resistant and grows at 0.5 µg/mL (**Fig. 41a**), with a complete growth inhibition at 2 µg/mL, as previously reported (23, 46). Strikingly, the three *yej* mutants were even more resistant with a higher growth at 0.5 µg/mL (**Fig. 41a**) and a complete growth inhibition at 4 µg/mL, suggesting that the transporter contributes to bleomycin uptake.

The Bac7 peptide (its name is unrelated to BacA or bacteroids) is a ribosome-inhibiting AMP of animal origin that is imported via SbmA or BacA (47, 48). With a flow cytometry-based assay and a fluorescent derivative of the Bac7 peptide, we found that the three *yej* mutants

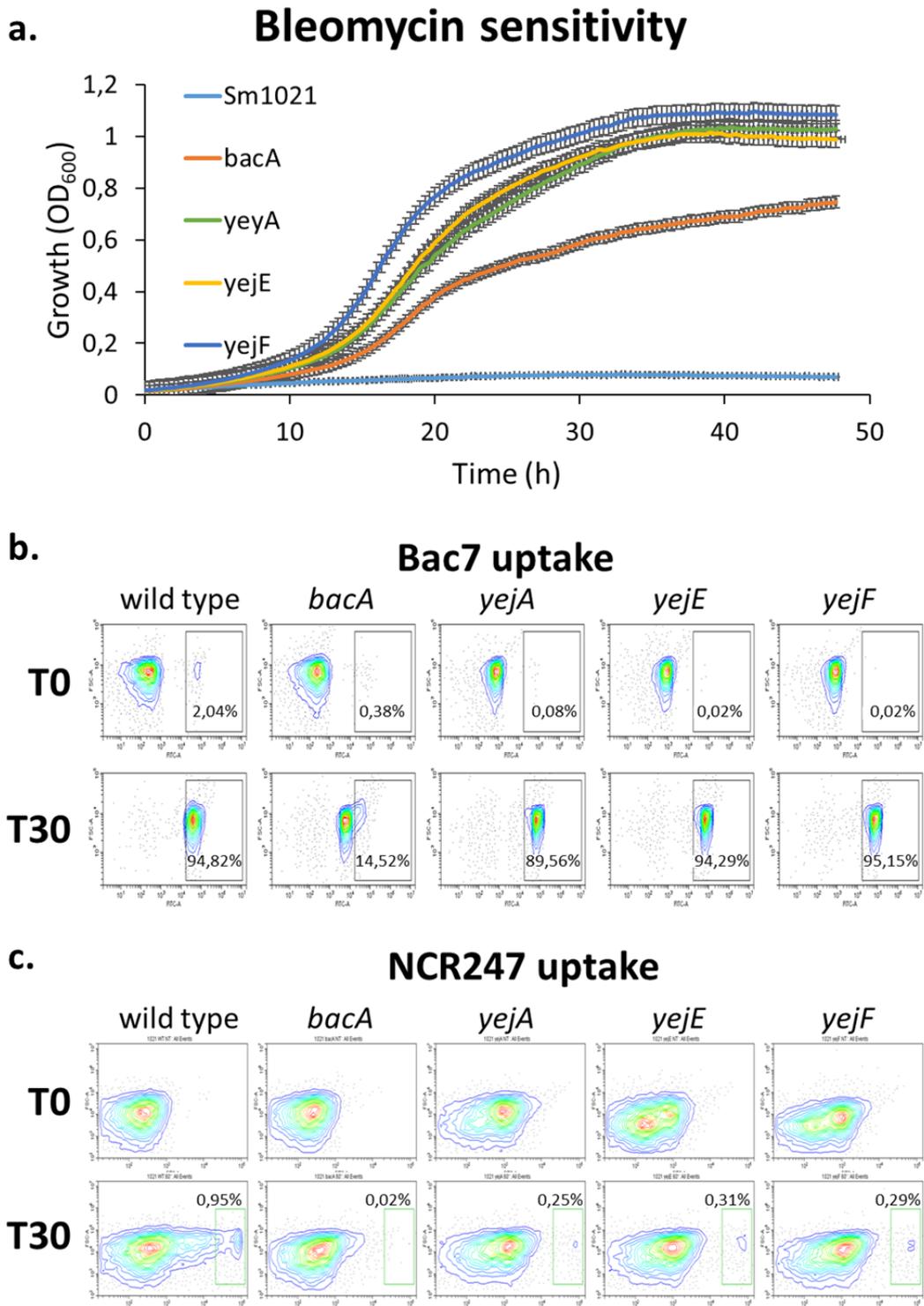


Figure 41 | Uptake of the Bac7 and NCR247 peptide mediated by YejABEF. a | Bleomycin uptake. Sensitivity to bleomycin is used as an estimation of the uptake capacity of the tested strains. It is determined by measurement of growth in the presence of 0.5 $\mu\text{g}/\text{mL}$ bleomycin. **b |** Bac7 uptake. Flow cytometry assay of the uptake of the Bac7-FITC peptide. The wild type and *yej* mutants become fluorescent after incubation with the Bac7-FITC peptide indicating internalization of the peptide. The *bacA* mutant is unable to import the peptide. **c |** NCR247 uptake. Flow cytometry assay of the NCR247-FITC peptide. In the wild type, a subpopulation of the bacteria imports the NCR247-FITC peptide and in the *bacA* mutant, this population is absent. In the *yeyA*, *yejE* and *yejF* mutants, the FITC-positive subpopulation is strongly reduced compared to wild type.

were not affected in the internalization of Bac7 (**Fig. 41b**), excluding the involvement of the YejABEF transporter in the uptake of this peptide, contrary to BacA.

Finally, the NCR247 peptide is another known substrate of BacA (23, 24). We tested the impact of the YejABEF transporter on NCR247 uptake with a similar assay as for the Bac7 test above (**Fig. 41c**). The analysis showed that while the NCR247 uptake is completely abolished in the *bacA* mutant, as expected, and shown before, its uptake is also reduced but not completely abolished in the three *yej* mutants. This suggests that the transporter contributes to NCR uptake.

vi. Discussion

Multiple functions of Sinorhizobium meliloti contribute to NCR resistance and are required for bacteroid formation and persistence

The NCR peptides are a two-edged sword. On the one hand, they maneuver the rhizobial endosymbionts into a terminally differentiated state via a multitude of activities on the bacteria from which the membrane permeabilization, cell cycle perturbation (polyploidization) and cell enlargement are the most visible ones. On the other hand, NCRs are antimicrobial, potentially killing the endosymbionts. Therefore, rhizobia have to defend themselves to be able to establish a chronic infection in the NCR-producing symbiotic cells of the nodules. Previous work by us and by others has identified the BacA peptide transporter, barriers in the bacterial envelope constituted by the EPS and proteolytic degradation of the NCRs as defenses of *Sinorhizobium* strains against the NCRs of their *Medicago* hosts (14, 25, 27, 28). Moreover, in the *Bradyrhizobium* strains infecting the NCR-producing *Aeschynomene* legumes, other NCR-protective functions have been suggested besides the BacA-like transporter BclA, including the hopanoid outer membrane lipids (which are not produced by *Sinorhizobium* strains) as well as the peptidoglycan cell wall polymer (23, 24, 49–51).

Here, we defined three new functions in *S. meliloti*, the LPS, the sigma factor RpoH1, and the YejABEF peptide transporter as additional determinants in bacteroids, required to cope with the NCR peptides. We show that knockout mutants in these genes are more sensitive to a panel of antimicrobial NCRs and that this hypersensitivity is correlated with a strongly

enhanced membrane permeability of the nodule bacteria and abnormalities in their morphology and ploidy levels. It is striking, however, that the different mutants have markedly different bacteroid phenotypes, ranging from undifferentiated to hyper-differentiated. This can be attributed to at least three factors. Each mutant has a specific NCR-sensitivity profile when tested against a small panel of peptides and the “NCR landscape” present in the developing symbiotic cells is continuously changing because of the expression of the hundreds of NCR genes in different waves during symbiotic cell differentiation (19). Accordingly, each mutant could accumulate NCR-induced damage at a different rate and reach the breaking point at different stages over the course of the bacteroid differentiation process (**Supp. Fig. B1**). Moreover, the expression pattern of the bacterial genes in the nodule zones is different, suggesting that their principal impact is situated at different stages of the symbiotic cell development and bacteroid differentiation. In addition, we cannot exclude that these genes provide supplementary contributions to the survival of the nodule bacteria, other than the response to the NCR peptides. Nevertheless, blocking the targeting of the NCR peptides to the nodule bacteria by the use of the *M. truncatula dnf1* mutant prevents the membrane permeabilization of the nodule bacteria in the mutants (although the bacteroid differentiation is not happening in this plant mutant, precisely because of the absence of the NCR peptides). This observation places the symbiotic role of these bacterial genes downstream of the peptide targeting to the symbiosomes (except for *lpsB*, see below) and links them, at least in part, to the NCRs.

The functions of *S. meliloti* described here and before probably picture only part of the full toolkit of this symbiont to survive the NCR challenge. Many additional *S. meliloti* mutants are described with symbiotic phenotypes that are suggestive for a similar contribution to NCR resistance. Interesting candidate functions are the two-component regulators ExoS-ChvI and FeuP-FeuQ, which are upregulated by NCR treatment (31) and which are essential for symbiosis (52, 53), the cellular redox regulation via the *grx1*, *trxA*, *trxB*, *trxC*, *gshA*, *gshB*, *gor*, *sodA*, *katB* and *katC* genes, also required for symbiosis and bacteroid formation or persistence (54–59), the LysR-type transcriptional regulator LsrB (60), the protease complexes ClpXP and HslUV and GroEL chaperones, involved in protein quality control and turnover (21, 61, 62), and the outer membrane transport protein TolC (63). In addition to these genes that were first described in symbiosis in a context unrelated to the NCR peptides, other genes that could be important for

symbiosis, were discovered in a Tn-seq screen with the peptide NCR247 (30). This screen identified here described functions like the BacA and YejABEF transporters, the LPS and EPS biosynthesis, and the FeuP-FeuQ two-component regulator but also never analyzed functions including several peptidoglycan modifying enzymes, which remind the peptidoglycan modifying enzymes previously identified in bradyrhizobia (23, 51), as well as several transcription factors and ribosome-associated proteins. It would be of great interest to explore the function of these genes in relation to the NCR response of bacteroids. Note that some of these genes might have escaped identification in previous genetic screens for symbiosis mutants because mutations might provoke only very subtle phenotypes with a weak or even imperceptible overall effect on nitrogen fixation in standard laboratory conditions but with bacteroid alterations in morphology and persistence as illustrated here with the *yejABEF* mutants.

Response to NCR-induced stress regulated by the alternative sigma factor RpoH1

Bacteria deal with different types of stress conditions by global transcriptional responses mediated by alternative sigma factors that are specialized RNA polymerase subunits, recognizing subsets of promoters of stress-responsive genes. As a general classification, the RpoH sigma factors are believed to control cytoplasmic stress responses, while the RpoE sigma factors are known to respond to periplasmic and membrane stress (64). Given the membrane damage provoked by NCRs on the one hand, and the general function in bacteria of RpoE sigma factors in periplasmic and membrane stress response, including in response to AMPs (65), a role of the RpoE regulators of *S. meliloti* in bacteroid differentiation and NCR response could a priori have been expected. *S. meliloti* has 11 RpoE-like sigma factors. Remarkably, despite the fact that some RpoEs have a considerable effect on gene transcription, all single mutants, all possible double mutants and even a mutant lacking all 11 genes showed no detectable phenotypic difference with wild type in symbiosis or during many tested free-living growth conditions, including growth in the presence of various stresses amongst which were membrane stresses (66, 67). Thus, the 11 *S. meliloti* RpoEs do not have the expected role in regulating the envelope stress response.

This role seems to be taken up in part by RpoH in *S. meliloti*, in agreement with its role in symbiosis and NCR resistance. The *S. meliloti* RpoH1 sigma factor controls the expression of

osmotic, oxidative and pH stress response genes (38) and growth of the *rpoH1* mutant is affected at low pH, high temperature as well as in the presence of various membrane disrupting agents (40). The latter suggests that RpoH in *S. meliloti* is not respecting the canonical division of labor between alternative sigma factors and that its function is not restricted to cytoplasmic stress but also encompasses membrane stresses. We propose that the here uncovered role of RpoH1 in NCR resistance and in bacteroid differentiation is connected to its regulation of the oxidative stress and/or membrane stress.

It should be noted that in *Rhizobium etli* the RpoH sigma factors are also required for symbiosis with the legume *Phaseolus vulgaris* (68) despite the fact that this legume does not produce NCR peptides and that its bacteroids are not terminally differentiated (10, 12). This suggests that the function of RpoH in symbiosis is broader than the response to the NCR peptides.

The lipopolysaccharide barrier against NCR membrane damage

In gram-negative bacteria like rhizobia, LPS constitutes the outermost shield of the cell and thus the first point of attack of AMPs, in particular the cationic AMPs. This attack is a two-stage process. The AMPs make first electrostatic interactions with the negatively charged LPS, allowing the AMPs to approach the membrane lipids and subsequently to insert into the lipid bilayer, perforate it and translocate into the periplasm (69). LPS is composed of three distinct domains: the lipid A moiety, anchoring the LPS molecules in the outer leaflet of the outer membrane; the core oligosaccharide region attached to the lipid A; and the most external O-antigen polysaccharide, attached to the lipid A. The chemical composition of each of these domains can influence the efficiency of the AMP attack and pathogens have evolved mechanisms to recognize the presence of AMPs and to modify in response the composition of the LPS to lower the potency of the AMP attack (70).

S. meliloti *lpsB* encodes a glycosyl transferase that participates in the biosynthesis of the LPS core (71). Although the chemical structure of the LPS core in *S. meliloti* has not been determined yet, the *lpsB* mutation strongly affects the monosaccharide composition of the core (43). Despite the aberrant core region, LPS with O-antigen is still produced by the mutant but only with a short polymer while the wild type strain produces LPS with short and long O-

antigens (42). The O-antigen is believed to constitute a camouflage, masking the membrane and the charges in the membrane vicinity. Thus, in the *lpsB* mutant the O-antigen could be a less efficient shield against the NCRs or alternatively, the changed core could offer to the peptides an easier access to the membrane-proximal charges on lipid A and the inner-core. Either both mechanisms or one of them could contribute to the increased sensitivity of the *lpsB* mutant to the NCRs (this work) as well as to other AMPs (36, 43).

The *lpsB* mutant forms nodules on *M. sativa*, which contain infected cells and has a phenotype that is quite similar to the *bacA* mutant. However, this *lpsB* mutant forms uninfected nodules on *M. truncatula* roots. Therefore, it was not possible to test the implication of NCRs in this symbiotic phenotype with the use of the *dnf1* mutant, since the *DNF1* gene is expressed only in infected nodule cells (44). Since also the NCR genes are nearly exclusively expressed in the infected symbiotic cells (19), the blockage of the mutant before release of bacteria into nodule cells suggests that the LPS could provide protection against another stressor produced very early on in the infection process in *M. truncatula*. On the other hand, it was recently reported that some NCR genes are expressed in infected root hairs or Nod factor stimulated root epidermal cells (72, 73). Thus, it is possible that the challenge with these early NCR peptides is already detrimental to the mutant, blocking any further progress in the infection process.

LpxXL is a specific acyl transferase that introduces in the lipid A moiety of LPS the very long chain fatty acid 27-OHC28:0 (41). This LpxXL-dependent acylation is expected to make the lipid A hydrophobic, forming the biophysical basis for LpxXL-dependent NCR resistance. Hydrophobic lipid A increases the thickness of the outer layer of the outer membrane and reduces the membrane fluidity, which on its turn prevents or delays AMP insertion and membrane damage. Increasing lipid A hydrophobicity by introducing additional acyl chains is a well-known mechanism used by *Salmonella* to enhance its resistance against host AMPs during infection (70). The *S. meliloti* acyl carrier protein AcpXL functions in conjunction with LpxXL to introduce the very long chain fatty acid on lipid A (37, 41). The *acpXL* mutant has similar phenotypes in our hands as *lpxXL* in terms of symbiosis and NCR sensitivity (data not shown). Furthermore, a set of other *S. meliloti* mutants in LPS biosynthesis, mostly with less defined biochemical functions, have symbiotic defects and are more sensitive to AMPs (36). It

will of interest to measure in detail the resistance of these mutants to NCR peptides and to analyze with histological methods the nodule formation and bacteroid differentiation by them.

The YejABEF peptide transporter provides resistance to membrane-damaging peptides and sensitizes bacteria to AMPs with intracellular targets

The YejABEF ABC transporter has never been analyzed before in the context of the rhizobium-legume symbiosis. Furthermore, even though this transporter is highly conserved among proteobacteria, its physiological role in whichever bacterium has been characterized only in a few instances. One of them is the uptake in *E. coli* of microcin C, a translation-inhibiting AMP of bacterial origin. This peptide has no action on the bacterial membrane but has an intracellular target. A mutant in the YejABEF transporter cannot uptake microcin C and is resistant to the AMP (32, 74). Conversely, the *Salmonella* and *Brucella* mutants are more sensitive to peptides with membrane damaging activity such as defensins, polymyxin B, protamine, melittin. Consequently, these mutants have reduced pathogenicity because their capacity to survive in macrophages and in animals is reduced (33, 34). Thus, the increased sensitivity of the *yejA*, *yejE* and *yejF* mutants of *S. meliloti* towards NCRs, which was suggested by Tn-seq (30) and which we demonstrate here, corresponds with these previous findings.

The characteristics of the YejABEF peptide transporter towards membrane-damaging peptides versus peptides with intracellular action are intriguingly parallel to the features of the BacA peptide transporter (SbmA in *E. coli* and *Salmonella*). These latter transporters are required for the import of diverse peptides with intracellular targets (microcins B and J, Bac7, Bac5, bleomycin) and mutants are therefore resistant to them while the same mutants are hypersensitive to membrane-active peptides (defensins, NCRs) (25, 46, 75). Moreover, the range of peptides that can be imported overlap between these two types of transporters. Both, YejABEF and BacA (or SbmA) can uptake bleomycin and extended derivatives of microcin C (this work; 32). Furthermore, we show that both BacA and YejABEF contribute to NCR247 uptake. Intriguingly, our genetic analysis suggests that both transporters cooperate to import this peptide, since inactivation of one or the other abolishes or strongly reduces uptake. How these two transporters can physically interact and cooperate is an issue of interest for future research. Nevertheless, BacA can function (partially) in the absence of YejABEF while the inverse is not the case. This could be the basis for the markedly different symbiotic phenotype of the

bacA and *yej* mutants. While the *bacA* mutation is detrimental from the earliest contact with NCRs and bacteria die as soon as they are released from infection threads in the symbiotic cells, the *yej* mutants cope with the NCR peptides much longer and show abnormalities only at the end of the bacteroid differentiation process. The *yejA* gene encodes the periplasmic binding protein of the transporter. The *yejA* mutant had a different NCR sensitivity profile than the *yejE* and *yejF* mutant and this was correlated with a different symbiotic phenotype in the *M. sativa* host. This suggests that *YejA* contributes to the interaction of only a subset of the NCR peptides.

How does the YejABEF and BacA transporters contribute to resistance to NCRs and other membrane-damaging peptides? The most straightforward model that has been proposed before for BacA (23–25) as well as for the unrelated SapABCDF peptide uptake transporter (76), is the reduction below a critical threshold of the AMP concentration in the vicinity of the inner membrane. Alternatively, the presence or activity of the transporters might indirectly affect the bacterial envelope structure, rendering it more robust against AMPs. The higher sensitivity of the *yej* mutants towards SDS is in agreement with this possibility. A similar detergent sensitivity and membrane alterations were reported in the *bacA* mutant of *S. meliloti* (77). In that respect, we note that the abnormal bacteroid morphology and increased ploidy level of the *yejE* and *yejF* mutants shows striking similarities with the DD-CPase1 peptidoglycan synthesis mutant in *Bradyrhizobium* sp. ORS285 in interaction with *Aeschynomene afraspera* (23, 51). It is possible that the YejABEF transporter function affects peptidoglycan structure by a still unknown mechanism.

Finally, in order to enforce the link between the symbiotic function of the YejABEF transporter and the presence of NCR peptides in nodules, we believe that it would be of interest to analyze mutants in the homologous genes of other rhizobia (e.g. *bradyrhizobia* interacting with NCR-producing *Aeschynomene* or rhizobia interacting with NCR-lacking legumes like soybean).

vii. Conclusions

The multifaceted NCR resistance required for symbiosis and chronic infection of the nodule cells mirrors the multitude of AMP resistance mechanisms in animal pathogens, which

collectively contribute to the pathogenicity of these bacteria (2, 3, 78). However, one dimension of this strategy in pathogens is not known in *S. meliloti* and consists of the direct recognition of host AMPs by receptors triggering an adaptive response. Probably the best studied AMP receptor is the two-component regulator PhoPQ in *Salmonella*, which adjusts the lipid A, core oligosaccharide and O-antigen composition in response to the presence of host peptides (70). Perhaps *S. meliloti* ExoS-ChvI or FeuP-FeuQ is such a regulatory module, recognizing NCRs and controlling an appropriate response in the symbiotic nodule cells? In this respect, it is of interest to note that the rhizobial LPS structure changes strongly in bacteroids of NCR-producing nodules (79, 80).

As shown here, rhizobia have to defend themselves to be able to establish a chronic infection in the NCR-producing symbiotic cells of the nodules. On the other hand, the profile of NCR peptides produced in the nodule cells is also determinant for the outcome of the symbiosis and some *M. truncatula* mutants in individual NCR genes or *M. truncatula* accessions, expressing specific NCR alleles, display incompatibility with *S. meliloti* strains (81–84). Thus, a fine balance must be established in the symbiotic nodule cells between the NCR landscapes and matching multifactorial bacterial countermeasures. Perturbations in the host or in the endosymbiont, as the ones described here, affecting this equilibrium leads to a breakdown of the symbiosis.

viii. Materials and Methods

Bacterial strains, plant growth and nodulation assays and analysis

The procedures for the growth of the *S. meliloti* strain Sm1021 and its derivatives, plant culture of the *M. sativa* cultivar *Gabès*, the *M. truncatula* accession *Jemalong A17*, and the *A17 dnf1* mutant, nodulation assays, acetylene reduction assays, bacteroid isolation, flow cytometry measurements and confocal microscopy were performed as described before (85). Bacterial mutants in *rpoH1*, *yejA*, *yejE* and *yejF* were obtained by plasmid insertion and gene deletion (**Table S2**) as described before (51).

In vitro sensitivity and peptide uptake assays

NCR and bleomycin sensitivity assays were carried out essentially as described (23). To measure the resistance of strains to sodium dodecyl sulfate (SDS), H₂O₂, HCl and NaCl stress, overnight cultures of the wild type and mutants were diluted to OD₆₀₀=0.2. A total of 100 µL of these suspensions was added to 3 ml soft agar (0.7% agar) and poured onto standard 1.5% agar plates. After solidification of the soft agar, filter paper disks (5 mm diameter) were placed on the center of the plate, and 5 µl of 10% (w/v) SDS, 2 M HCl or 30% H₂O₂ was added to the disks. Plates were incubated at 30°C for 3 days, and the diameter of the clearing zone was measured. Peptide uptake assays were performed as described (23, 86).

TEM

Bacterial suspensions at OD₆₀₀=6 or 21 days old nodule samples were incubated in fixative (3% glutaraldehyde, 1% paraformaldehyde in 0.1M cacodylate pH 6.8) for one hour, and washed with cacodylate buffer 0.1 M, pH 6.8. Samples were then incubated in 1% osmium tetroxide, 1.5% potassium ferrocyanide in water. After washing, samples were dehydrated by incubation in increasing concentrations of ethanol (10-20-30-50-70-90-100%-absolute ethanol-propylene oxide), followed by infiltration with epoxy resin (low viscosity Premix Kit medium, Agar) and polymerization for 16h at 60°C. Ultrathin sections (80-70 nm) were obtained with an ultramicrotome EM UC6 (Leica Microsystems) and collected on formvar carbon-coated copper grids (Agar). Sections were stained with 2% uranyl acetate (Merck) and lead citrate (Agar) before observation with a JEOL JEM-1400 transmission electron microscope operating at 120kV. Images were acquired using a postcolumn high-resolution (11 megapixels) high-speed camera (SC1000 Orius; Gatan) and processed with Digital Micrograph (Gatan).

ix. Acknowledgements

We are grateful to Corinne Foucault for technical assistance. QN was supported by a PhD fellowship from the Paris-Saclay University and QB and NB by Post-doc grants from the Agence Nationale de la Recherche. The present work has benefited from Imagerie-Gif core facility supported by the Agence Nationale de la Recherche (ANR-11-EQPX-0029/Morphoscope, ANR-10-INBS-04/FranceBioImaging; ANR - 11 - IDEX - 0003 - 02/ Saclay Plant Sciences). This work

was supported by the LabEx Saclay Plant Sciences-SPS and the Agence Nationale de la Recherche, grants ANR-17-CE20-0011 and ANR-16-CE20-0013.

x. Author contributions

QN, QB and NB have initiated jointly this work. QN, QB, NB, ML, RL, CB and TT constructed strains and performed nodulation assays. QN, QB, RL, CB and TT performed microscopy and flow cytometry. QB, SJ, AK and EK performed the in vitro experiments and provided reagents. QN, QB, NB, TT, BA and PM conceived the study and analyzed the data. PM wrote the manuscript with input from all authors.

xi. References

1. Hanson MA, Lemaitre B. 2020. New insights on *Drosophila* antimicrobial peptide function in host defense and beyond. *Curr Opin Immunol* 62:22–30.
2. Mergaert P. 2018. Role of antimicrobial peptides in controlling symbiotic bacterial populations. *Nat Prod Rep* 35:336–356.
3. Maróti G, Kereszt A, Kondorosi E, Mergaert P. 2011. Natural roles of antimicrobial peptides in microbes, plants and animals. *Res Microbiol* 162:363–374.
4. Mergaert P, Kikuchi Y, Shigenobu S, Nowack ECM. 2017. Metabolic integration of bacterial endosymbionts through antimicrobial peptides. *Trends Microbiol* 25:703–712.
5. Kondorosi E, Mergaert P, Kereszt A. 2013. A paradigm for endosymbiotic life: cell differentiation of *Rhizobium* bacteria provoked by host plant factors. *Annu Rev Microbiol* 67:611–628.
6. Alunni B, Gourion B. 2016. Terminal bacteroid differentiation in the legume-rhizobium symbiosis: nodule-specific cysteine-rich peptides and beyond. *New Phytol* 211:411–417.
7. Pan H, Wang D. 2017. Nodule cysteine-rich peptides maintain a working balance during nitrogen-fixing symbiosis. *Nat Plants* 3:17048.
8. Stonoha-Arther C, Wang D. 2018. Tough love: accommodating intracellular bacteria through directed secretion of antimicrobial peptides during the nitrogen-fixing symbiosis. *Curr Opin Plant Biol* 44:155–163.
9. Roy P, Achom M, Wilkinson H, Lagunas B, Gifford ML. 2020. Symbiotic outcome modified by the diversification from 7 to over 700 nodule-specific cysteine-rich peptides. *Genes* 11:348.
10. Mergaert P, Uchiumi T, Alunni B, Evanno G, Cheron A, Catrice O, Mausset A-E, Barloy-Hubler F, Galibert F, Kondorosi A, Kondorosi E. 2006. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. *Proc Natl Acad Sci USA* 103:5230–5235.
11. 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, 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* 169:1254–1265.
12. Mergaert P, Nikovics K, Kelemen Z, Maunoury N, Vaubert D, Kondorosi A, 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.
13. Van de Velde W, Zehirov G, Szatmari A, Debreczeny M, Ishihara H, Kevei Z, Farkas A, Mikulass K, Nagy A, Tiricz H, Satiat-Jeunemaître B, Alunni B, Bourge M, Kucho K, Abe M, Kereszt A, Maroti G, Uchiumi T, Kondorosi E, Mergaert P. 2010. Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327:1122–1126.
14. Montiel J, Downie JA, Farkas A, Bihari P, Herczeg R, Bálint B, Mergaert P, Kereszt A, Kondorosi É. 2017. Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proc Natl Acad Sci USA* 114:5041–5046.

15. Alunni B, Kevei Z, Redondo-Nieto M, Kondorosi A, Mergaert P, Kondorosi E. 2007. Genomic organization and evolutionary insights on GRP and NCR genes, two large nodule-specific gene families in *Medicago truncatula*. *Mol Plant-Microbe Interact* 20:1138–1148.
16. Nallu S, Silverstein KAT, Zhou P, Young ND, Vandenbosch KA. 2014. Patterns of divergence of a large family of nodule cysteine-rich peptides in accessions of *Medicago truncatula*. *Plant J* 78:697–705.
17. Pecrix Y, Staton SE, Sallet E, Lelandais-Brière C, Moreau S, Carrère S, Blein T, Jardinaud M-F, Latrasse D, Zouine M, Zahm M, Kreplak J, Mayjonade B, Satgé C, Perez M, Cauet S, Marande W, Chantry-Darmon C, Lopez-Roques C, Bouchez O, Bérard A, Debellé F, Muños S, Bendahmane A, Bergès H, Niebel A, Buitink J, Frugier F, Benhamed M, Crespi M, Gouzy J, Gamas P. 2018. Whole-genome landscape of *Medicago truncatula* symbiotic genes. *Nat Plants* 4:1017–1025.
18. Gully D, Czernic P, Cruveiller S, Mahé F, Longin C, Vallenet D, François P, Nidelet S, Rialle S, Giraud E, Arrighi J-F, DasGupta M, Cartieaux F. 2018. Transcriptome profiles of Nod factor-independent symbiosis in the tropical legume *Aeschynomene evenia*. *Sci Rep* 8:10934.
19. Guefrachi I, Nagymihaly M, Pislariu CI, Van de Velde W, Ratet P, Mars M, Udvardi MK, Kondorosi E, Mergaert P, Alunni B. 2014. Extreme specificity of *NCR* gene expression in *Medicago truncatula*. *BMC Genom* 15:712.
20. Lima RM, Kylarová S, Mergaert P, Kondorosi É. 2020. Unexplored arsenals of legume peptides with potential for their applications in medicine and agriculture. *Front Microbiol* 11:1307.
21. Farkas A, Maróti G, Durgó H, Györgypál Z, Lima RM, Medzihradzsky KF, Kereszt A, Mergaert P, Kondorosi É. 2014. *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proc Natl Acad Sci USA* 111:5183–5188.
22. Cullen TW, Schofield WB, Barry NA, Putnam EE, Rundell EA, Trent MS, Degnan PH, Booth CJ, Yu H, Goodman AL. 2015. Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation. *Science* 347:170–175.
23. Barrière Q, Guefrachi I, Gully D, Lamouche F, Pierre O, Fardoux J, Chaintreuil C, Alunni B, Timchenko T, Giraud E, Mergaert P. 2017. Integrated roles of BclA and DD-carboxypeptidase 1 in Bradyrhizobium differentiation within NCR-producing and NCR-lacking root nodules. *Sci Rep* 7:9063.
24. Guefrachi I, Pierre O, Timchenko T, Alunni B, Barrière Q, Czernic P, Villaécija-Aguilar J-A, Verly C, Bourge M, Fardoux J, Mars M, Kondorosi E, Giraud E, Mergaert P. 2015. *Bradyrhizobium* BclA is a peptide transporter required for bacterial differentiation in symbiosis with *Aeschynomene* legumes. *Mol Plant-Microbe Interact* 28:1155–1166.
25. Haag AF, Baloban M, Sani M, Kerscher B, Pierre O, Farkas A, Longhi R, Boncompagni E, Hérouart D, Dall'angelo S, Kondorosi E, Zanda M, Mergaert P, Ferguson GP. 2011. Protection of *Sinorhizobium* against host cysteine-rich antimicrobial peptides is critical for symbiosis. *PLoS Biol* 9:e1001169.
26. Rempel S, Gati C, Nijland M, Thangaratnarajah C, Karyolaimos A, de Gier JW, Guskov A, Slotboom DJ. 2020. A mycobacterial ABC transporter mediates the uptake of hydrophilic compounds. *Nature* 580:409–412.
27. Arnold MFF, Penterman J, Shabab M, Chen EJ, Walker GC. 2018. Important late-stage symbiotic role of the *Sinorhizobium meliloti* exopolysaccharide succinoglycan. *J Bacteriol* 200:e00665-17.

28. Price PA, Tanner HR, Dillon BA, Shabab M, Walker GC, Griffiths JS. 2015. Rhizobial peptidase HrrP cleaves host-encoded signaling peptides and mediates symbiotic compatibility. *Proc Natl Acad Sci USA* 112:15244–15249.
29. Cole JN, Nizet V. 2016. Bacterial evasion of host antimicrobial peptide defenses. *Microbiol Spectr* 4.
30. Arnold MFF, Shabab M, Penterman J, Boehme KL, Griffiths JS, Walker GC. 2017. Genome-wide sensitivity analysis of the microsymbiont *Sinorhizobium meliloti* to symbiotically important, defensin-like host peptides. *mBio* 8:e01060-17.
31. Penterman J, Abo RP, De Nisco NJ, Arnold MFF, Longhi R, Zanda M, Walker GC. 2014. Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. *Proc Natl Acad Sci USA* 111:3561–3566.
32. Bantysh O, Serebryakova M, Zukher I, Kulikovskiy A, Tsibulskaya D, Dubiley S, Severinov K. 2015. Enzymatic synthesis and functional characterization of bioactive Microcin C-like compounds with altered peptide sequence and length. *J Bacteriol* 197:3133–3141.
33. Eswarappa SM, Panguluri KK, Hensel M, Chakravorty D. 2008. The yejABEF operon of *Salmonella* confers resistance to antimicrobial peptides and contributes to its virulence. *Microbiology* 154:666–678.
34. Wang Z, Bie P, Cheng J, Lu L, Cui B, Wu Q. 2016. The ABC transporter YejABEF is required for resistance to antimicrobial peptides and the virulence of *Brucella melitensis*. *Sci Rep* 6:31876.
35. Simpson BW, Trent MS. 2019. Pushing the envelope: LPS modifications and their consequences. *Nat Rev Microbiol* 17:403–416.
36. Campbell GRO, Sharypova LA, Scheidle H, Jones KM, Niehaus K, Becker A, Walker GC. 2003. Striking complexity of lipopolysaccharide defects in a collection of *Sinorhizobium meliloti* mutants. *J Bacteriol* 185:3853–3862.
37. Haag AF, Wehmeier S, Beck S, Marlow VL, Fletcher V, James EK, Ferguson GP. 2009. The *Sinorhizobium meliloti* LpxXL and AcpXL proteins play important roles in bacteroid development within alfalfa. *J Bacteriol* 191:4681–4686.
38. Barnett MJ, Bittner AN, Toman CJ, Oke V, Long SR. 2012. Dual RpoH sigma factors and transcriptional plasticity in a symbiotic bacterium. *J Bacteriol* 194:4983–4994.
39. Tiricz H, Szucs A, Farkas A, Pap B, Lima RM, Maróti G, Kondorosi É, Kereszt A. 2013. Antimicrobial nodule-specific cysteine-rich peptides induce membrane depolarization-associated changes in the transcriptome of *Sinorhizobium meliloti*. *Appl Environ Microbiol* 79:6737–6746.
40. Mitsui H, Sato T, Sato Y, Ito N, Minamisawa K. 2004. *Sinorhizobium meliloti* RpoH1 is required for effective nitrogen-fixing symbiosis with alfalfa. *Mol Genet Genom* 271:416–425.
41. Ferguson GP, Datta A, Carlson RW, Walker GC. 2005. Importance of unusually modified lipid A in *Sinorhizobium* stress resistance and legume symbiosis. *Mol Microbiol* 56:68–80.
42. Niehaus K, Lagares A, Pühler A. 1998. A *Sinorhizobium meliloti* lipopolysaccharide mutant induces effective nodules on the host plant *Medicago sativa* (alfalfa) but fails to establish a symbiosis with *Medicago truncatula*. *Mol Plant-Microbe Interact* 11:906–914.

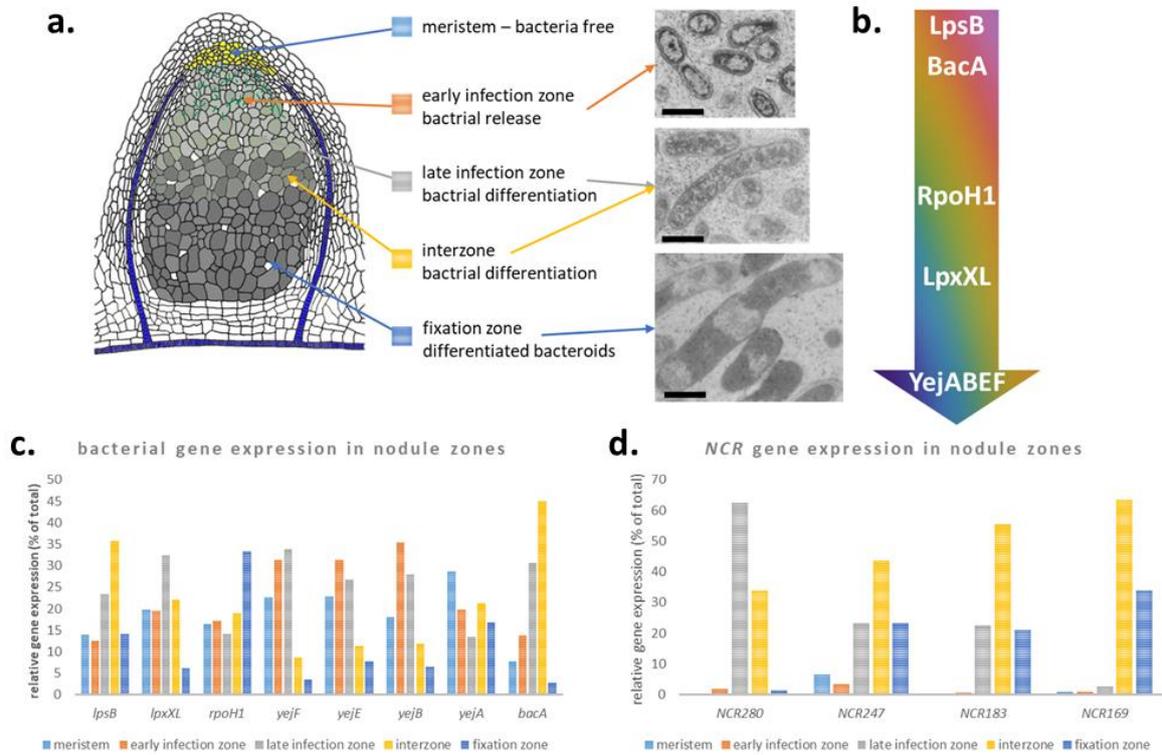
43. Campbell GRO, Reuhs BL, Walker GC. 2002. Chronic intracellular infection of alfalfa nodules by *Sinorhizobium meliloti* requires correct lipopolysaccharide core. Proc Natl Acad Sci USA 99:3938–3943.
44. Wang D, Griffiths J, Starker C, Fedorova E, Limpens E, Ivanov S, Bisseling T, Long S. 2010. A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. Science 327:1126–1129.
45. Mendoza-Suárez MA, Geddes BA, Sánchez-Cañizares C, Ramírez-González RH, Kirchhelle C, Jorrián B, Poole PS. 2020. Optimizing *Rhizobium*-legume symbioses by simultaneous measurement of rhizobial competitiveness and N₂ fixation in nodules. Proc Natl Acad Sci USA 117:9822–9831.
46. Wehmeier S, Arnold MFF, Marlow VL, Aouida M, Myka KK, Fletcher V, Benincasa M, Scocchi M, Ramotar D, Ferguson GP. 2010. Internalization of a thiazole-modified peptide in *Sinorhizobium meliloti* occurs by BacA-dependent and -independent mechanisms. Microbiology 156:2702–2713.
47. Marlow VL, Haag AF, Kobayashi H, Fletcher V, Scocchi M, Walker GC, Ferguson GP. 2009. Essential role for the BacA protein in the uptake of a truncated eukaryotic peptide in *Sinorhizobium meliloti*. J Bacteriol 191:1519–1527.
48. Mardirossian M, Grzela R, Giglione C, Meinel T, Gennaro R, Mergaert P, Scocchi M. 2014. The host antimicrobial peptide Bac7₁₋₃₅ binds to bacterial ribosomal proteins and inhibits protein synthesis. Chem Biol 21:1639–1647.
49. Kulkarni G, Busset N, Molinaro A, Gargani D, Chaintreuil C, Silipo A, Giraud E, Newman DK. 2015. Specific hopanoid classes differentially affect free-living and symbiotic states of *Bradyrhizobium diazoefficiens*. mBio 6:e01251-01215.
50. Silipo A, Vitiello G, Gully D, Sturiale L, Chaintreuil C, Fardoux J, Gargani D, Lee H-I, Kulkarni G, Busset N, Marchetti R, Palmigiano A, Moll H, Engel R, Lanzetta R, Paduano L, Parrilli M, Chang W-S, Holst O, Newman DK, Garozzo D, D'Errico G, Giraud E, Molinaro A. 2014. Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. Nat Comm 5:5106.
51. Gully D, Gargani D, Bonaldi K, Grangeteau C, Chaintreuil C, Fardoux J, Nguyen P, Marchetti R, Nouwen N, Molinaro A, Mergaert P, Giraud E. 2016. A peptidoglycan-remodeling enzyme is critical for bacteroid differentiation in *Bradyrhizobium* spp. during legume symbiosis. Mol Plant-Microbe Interact 29:447–457.
52. Bélanger L, Dimmick KA, Fleming JS, Charles TC. 2009. Null mutations in *Sinorhizobium meliloti* *exoS* and *chvI* demonstrate the importance of this two-component regulatory system for symbiosis. Mol Microbiol 74:1223–1237.
53. Griffiths JS, Carlyon RE, Erickson JH, Moulton JL, Barnett MJ, Toman CJ, Long SR. 2008. A *Sinorhizobium meliloti* osmosensory two-component system required for cyclic glucan export and symbiosis. Mol Microbiol 69:479–490.
54. Santos R, Hérouart D, Puppo A, Touati D. 2000. Critical protective role of bacterial superoxide dismutase in rhizobium-legume symbiosis. Mol Microbiol 38:750–759.
55. Jamet A, Sigaud S, Van de Sype G, Puppo A, Hérouart D. 2003. Expression of the bacterial catalase genes during *Sinorhizobium meliloti*-*Medicago sativa* symbiosis and their crucial role during the infection process. Mol Plant-Microbe Interact 16:217–225.

56. Castro-Sowinski S, Matan O, Bonafede P, Okon Y. 2007. A thioredoxin of *Sinorhizobium meliloti* CE52G is required for melanin production and symbiotic nitrogen fixation. *Mol Plant-Microbe Interact* 20:986–993.
57. Benyamina SM, Baldacci-Cresp F, Couturier J, Chibani K, Hopkins J, Bekki A, de Lajudie P, Rouhier N, Jacquot J-P, Alloing G, Puppo A, Frendo P. 2013. Two *Sinorhizobium meliloti* glutaredoxins regulate iron metabolism and symbiotic bacteroid differentiation. *Environ Microbiol* 15:795–810.
58. Tang G, Li N, Liu Y, Yu L, Yan J, Luo L. 2018. *Sinorhizobium meliloti* glutathione reductase is required for both redox homeostasis and symbiosis. *Appl Environ Microbiol* 84:e01937-17.
59. Yang L, El Msehli S, Benyamina S, Lambert A, Hopkins J, Cazareth J, Pierre O, Hérouart D, Achi-Smiti S, Boncompagni E, Frendo P. 2020. Glutathione deficiency in *Sinorhizobium meliloti* does not impair bacteroid differentiation but induces early senescence in the interaction with *Medicago truncatula*. *Front Plant Sci* 11:137.
60. Tang G, Xing S, Wang S, Yu L, Li X, Staehelin C, Yang M, Luo L. 2017. Regulation of cysteine residues in LsrB proteins from *Sinorhizobium meliloti* under free-living and symbiotic oxidative stress. *Environ Microbiol* 19:5130–5145.
61. Bittner AN, Foltz A, Oke V. 2007. Only one of five *groEL* genes is required for viability and successful symbiosis in *Sinorhizobium meliloti*. *J Bacteriol* 189:1884–1889.
62. Ogden AJ, McAleer JM, Kahn ML. 2019. Characterization of the *Sinorhizobium meliloti* HslUV and ClpXP protease systems in free-living and symbiotic states. *J Bacteriol* 201:e00498-18.
63. Cosme AM, Becker A, Santos MR, Sharypova LA, Santos PM, Moreira LM. 2008. The outer membrane protein TolC from *Sinorhizobium meliloti* affects protein secretion, polysaccharide biosynthesis, antimicrobial resistance, and symbiosis. *Mol Plant-Microbe Interact* 21:947–957.
64. Gruber TM, Gross CA. 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* 57:441–466.
65. Guest RL, Raivio TL. 2016. Role of the gram-negative envelope stress response in the presence of antimicrobial agents. *Trends Microbiol* 24:377–390.
66. Sauviac L, Philippe H, Phok K, Bruand C. 2007. An extracytoplasmic function sigma factor acts as a general stress response regulator in *Sinorhizobium meliloti*. *J Bacteriol* 189:4204–4216.
67. Lang C, Barnett MJ, Fisher RF, Smith LS, Diodati ME, Long SR. 2018. Most *Sinorhizobium meliloti* extracytoplasmic function sigma factors control accessory functions. *mSphere* 3:e00454-18.
68. Martínez-Salazar JM, Sandoval-Calderón M, Guo X, Castillo-Ramírez S, Reyes A, Loza MG, Rivera J, Alvarado-Affantranger X, Sánchez F, González V, Dávila G, Ramírez-Romero MA. 2009. The *Rhizobium etli* RpoH1 and RpoH2 sigma factors are involved in different stress responses. *Microbiology* 155:386–397.
69. Zasloff M. 2002. Antimicrobial peptides of multicellular organisms. *Nature* 415:389–395.
70. Dalebroux ZD, Miller SI. 2014. *Salmonellae* PhoPQ regulation of the outer membrane to resist innate immunity. *Curr Opin Microbiol* 17:106–113.
71. Lagares A, Hozbor DF, Niehaus K, Otero AJ, Lorenzen J, Arnold W, Pühler A. 2001. Genetic characterization of a *Sinorhizobium meliloti* chromosomal region in lipopolysaccharide biosynthesis. *J Bacteriol* 183:1248–1258.

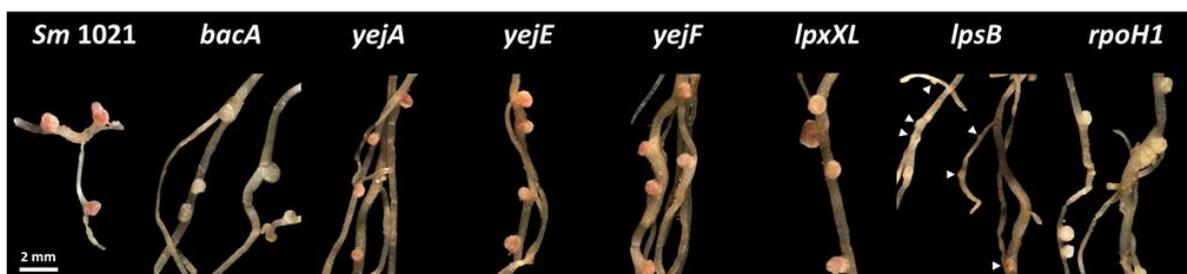
72. Jardinaud M-F, Boivin S, Rodde N, Catrice O, Kisiala A, Lepage A, Moreau S, Roux B, Cottret L, Sallet E, Brault M, Emery RJN, Gouzy J, Frugier F, Gamas P. 2016. A laser dissection-RNAseq analysis highlights the activation of cytokinin pathways by Nod factors in the *Medicago truncatula* root epidermis. *Plant Physiol* 171:2256–2276.
73. Liu C-W, Breakspear A, Guan D, Cerri MR, Jackson K, Jiang S, Robson F, Radhakrishnan GV, Roy S, Bone C, Stacey N, Rogers C, Trick M, Niebel A, Oldroyd GED, de Carvalho-Niebel F, Murray JD. 2019. NIN Acts as a network hub controlling a growth module required for rhizobial infection. *Plant Physiol* 179:1704–1722.
74. Novikova M, Metlitskaya A, Datsenko K, Kazakov T, Kazakov A, Wanner B, Severinov K. 2007. The *Escherichia coli* Yej transporter is required for the uptake of translation inhibitor microcin C. *J Bacteriol* 189:8361–8365.
75. Arnold MFF, Haag AF, Capewell S, Boshoff HI, James EK, McDonald R, Mair I, Mitchell AM, Kerscher B, Mitchell TJ, Mergaert P, Barry CE, Scocchi M, Zanda M, Campopiano DJ, Ferguson GP. 2013. Partial complementation of *Sinorhizobium meliloti* bacA mutant phenotypes by the *Mycobacterium tuberculosis* BacA protein. *J Bacteriol* 195:389–398.
76. Parra-Lopez C, Baer MT, Groisman EA. 1993. Molecular genetic analysis of a locus required for resistance to antimicrobial peptides in *Salmonella typhimurium*. *EMBO J* 12:4053–4062.
77. Ferguson GP, Roop RM, Walker GC. 2002. Deficiency of a *Sinorhizobium meliloti* BacA mutant in alfalfa symbiosis correlates with alteration of the cell envelope. *J Bacteriol* 184:5625–5632.
78. Bauer ME, Shafer WM. 2015. On the *in vivo* significance of bacterial resistance to antimicrobial peptides. *Biochim Biophys Acta* 1848:3101–3111.
79. Kannenberg EL, Carlson RW. 2001. Lipid A and O-chain modifications cause *Rhizobium* lipopolysaccharides to become hydrophobic during bacteroid development. *Mol Microbiol* 39:379–391.
80. Crespo-Rivas JC, Guefrachi I, Mok KC, Villaécija-Aguilar JA, Acosta-Jurado S, Pierre O, Ruiz-Sainz JE, Taga ME, Mergaert P, Vinardell JM. 2016. *Sinorhizobium fredii* HH103 bacteroids are not terminally differentiated and show altered O-antigen in nodules of the Inverted Repeat-Lacking Clade legume *Glycyrrhiza uralensis*. *Environ Microbiol* 18:2392–2404.
81. Kim M, Chen Y, Xi J, Waters C, Chen R, Wang D. 2015. An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proc Natl Acad Sci USA* 112:15238–15243.
82. Horváth B, Domonkos Á, Kereszt A, Szűcs A, Ábrahám E, Ayaydin F, Bóka K, Chen Y, Chen R, Murray JD, Udvardi MK, Kondorosi É, Kaló P. 2015. Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the *Medicago truncatula* dnf7 mutant. *Proc Natl Acad Sci USA* 112:15232–15237.
83. Wang Q, Yang S, Liu J, Terecskei K, Ábrahám E, Gombár A, Domonkos Á, Szűcs A, Körmöczi P, Wang T, Fodor L, Mao L, Fei Z, Kondorosi É, Kaló P, Kereszt A, Zhu H. 2017. Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. *Proc Natl Acad Sci USA* 114:6854–6859.
84. Yang S, Wang Q, Fedorova E, Liu J, Qin Q, Zheng Q, Price PA, Pan H, Wang D, Griffiths JS, Bisseling T, Zhu H. 2017. Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proc Natl Acad Sci USA* 114:6848–6853.

85. Kazmierczak T, Nagymihály M, Lamouche F, Barrière Q, Guefrachi I, Alunni B, Ouadghiri M, Ibijbjen J, Kondorosi É, Mergaert P, Gruber V. 2017. Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium* strains. *Mol Plant-Microbe Interact* 30:399–409.
86. Benincasa M, Barrière Q, Runti G, Pierre O, Bourge M, Scocchi M, Mergaert P. 2016. Single Cell Flow Cytometry Assay for Peptide Uptake by Bacteria. *Bio-protocol* 6:e2038–e2038.

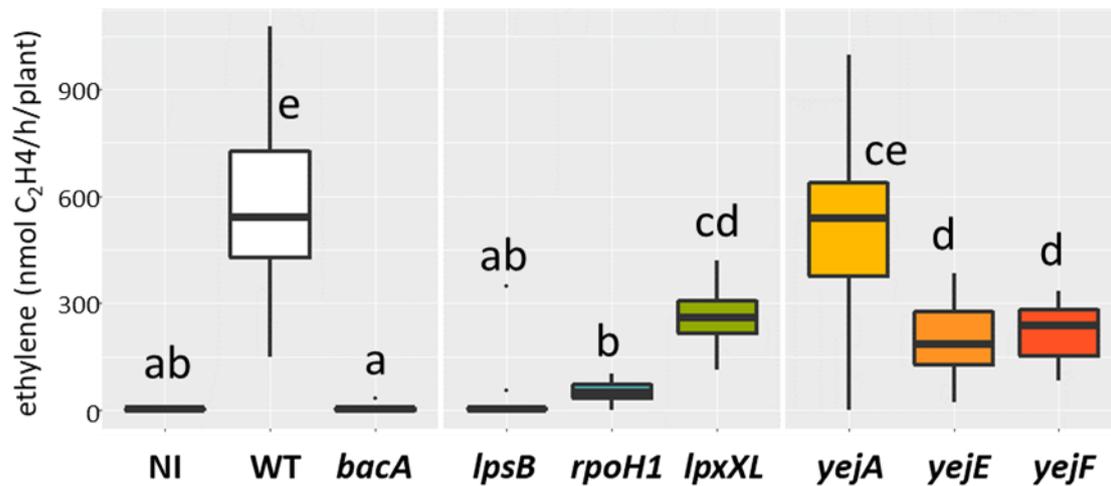
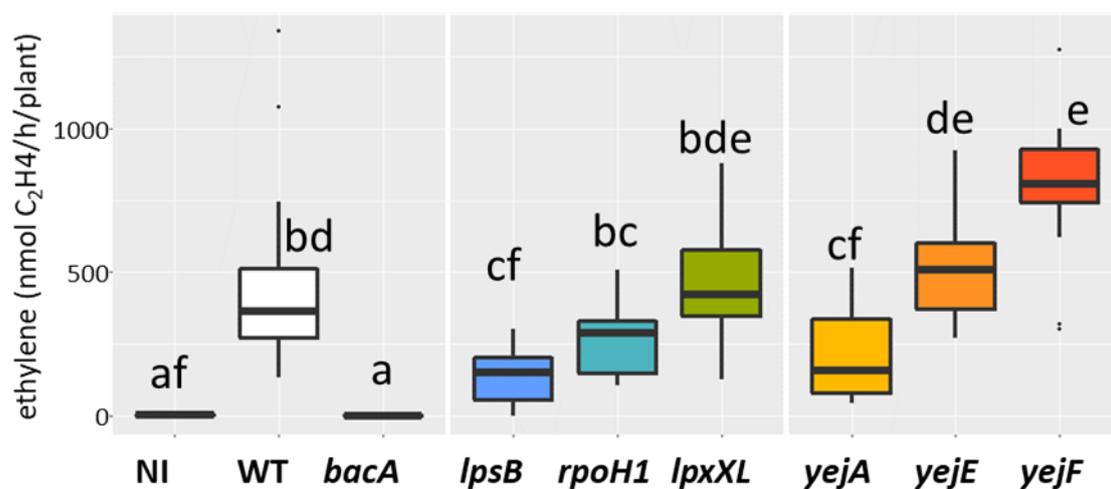
xii. Supplementary figures and tables



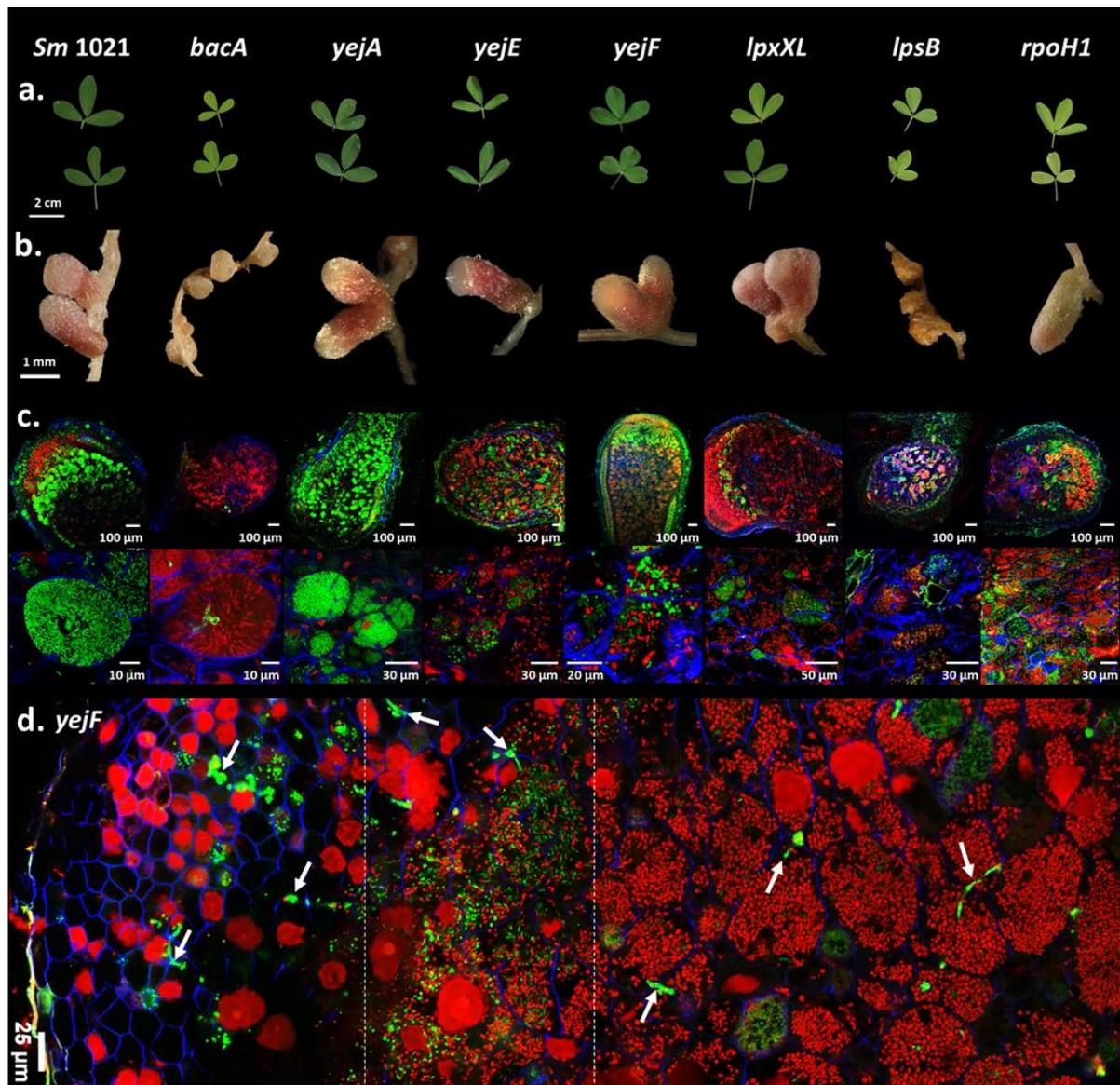
Supp. figure B1 | Expression pattern in nodules of bacterial and NCR genes. a | Schematic drawing of a *Medicago* nodule organized in tissues. The meristem is bacteria free and contains dividing cells allowing the organ to grow. In the early infection zone, bacteria proliferate in infection threads (green lines) and are released by endocytosis inside cells derived from the meristem. In the late infection zone and the so-called interzone, the bacteria differentiate into bacteroids. The fixation zone contains the fully differentiated, nitrogen-fixing bacteroids. The pictures show bacteria inside the nodule cells and are presented at the same scale (bar is 1 μ m) allowing to appreciate the transformation of the bacteria. **b |** Against the backdrop of a changing landscape of NCR peptides (rainbow colors representing peptides appearing and disappearing at different stages of symbiotic cell differentiation), the bacterial functions described in this study are critical at distinct stages of the bacteroid differentiation process. Based on the results of this study **c, d |** The relative expression profile (% of total) of the studied bacterial genes *lpsB*, *lpxXL*, *rpoH1*, *yejF*, *yejE*, *yejB*, *yejA* and *bacA* (**c**) and of *NCR280*, *NCR247*, *NCR183* and *NCR169* (**d**) in the meristem, early infection zone, late infection zone, interzone and fixation zone are displayed. Data was extracted from ref. 14 and was obtained by RNA-seq analysis on laser micro-dissected nodule tissues.



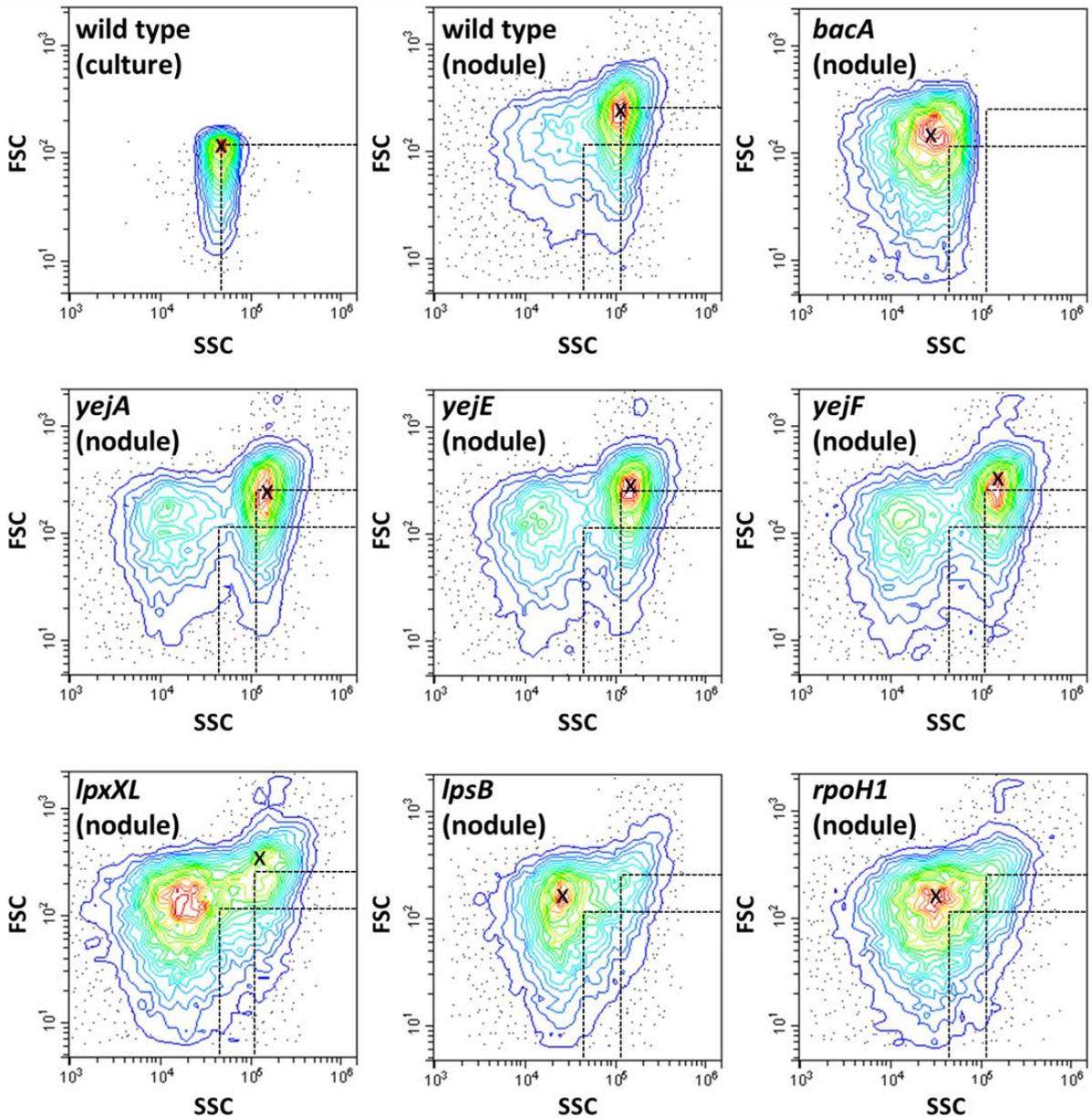
Supp. figure B2 | Nodule phenotypes at 21 days post inoculation. Arrowheads indicate small nodules elicited by the *lpsB* mutant. Scale bar = 2 mm.

a. *Medicago truncatula***b.** *Medicago sativa*

Supp. figure B3 | Nitrogen fixation activity determined by the acetylene reduction assay of nodulated plants infected with bacterial mutants. **a** | Acetylene reduction activity in *M. truncatula*. **b** | Acetylene reduction activity in *M. sativa*. Acetylene reduction assays were performed on whole roots of plants nodulated by the indicated mutants. NI, non-inoculated control plants; WT, plants nodulated by the wild-type strain Sm1021. Boxplots were generated from 15 plants each. Letters associated with each condition represent statistically different classes determined by a non-parametric Dunn test, with an α threshold equal to 0.05.

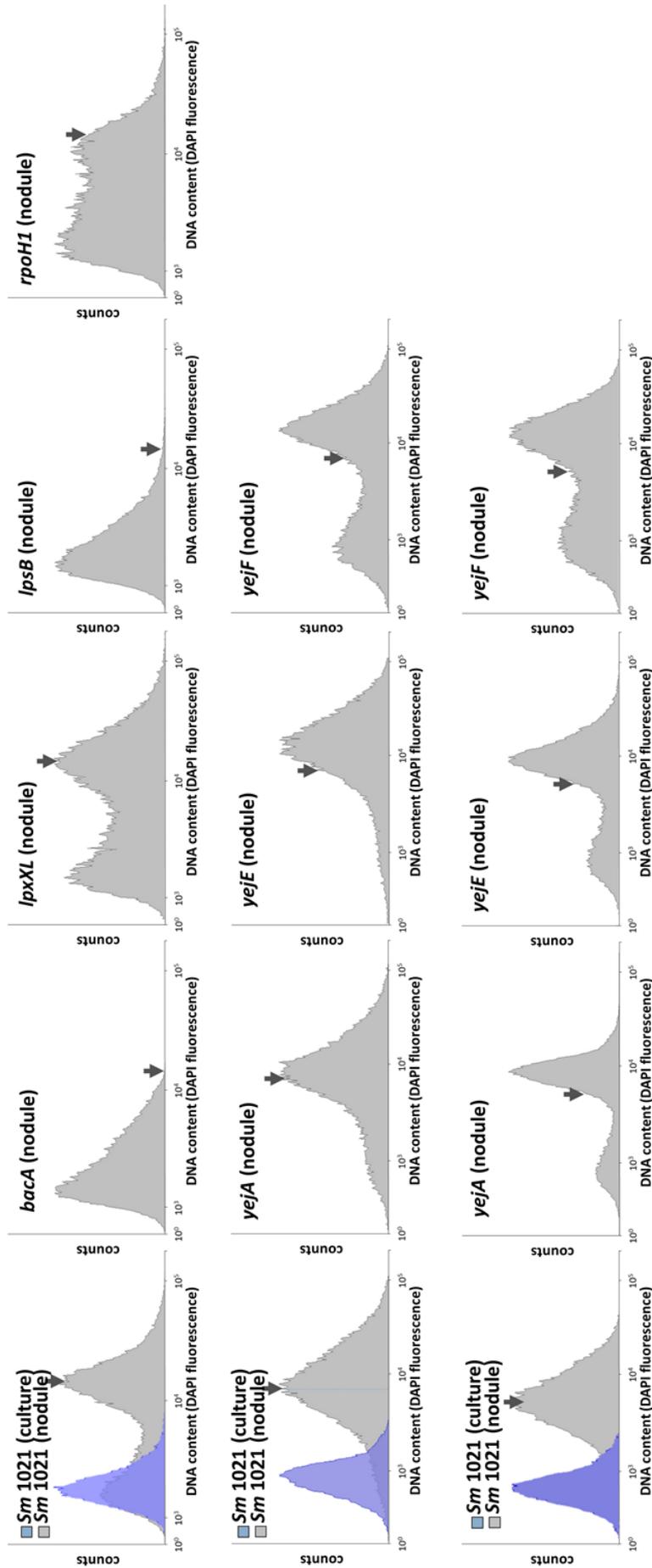


Supp. figure B4 | Symbiotic phenotype of *Sinorhizobium meliloti* 1021 mutants during symbiosis with *Medicago sativa*. **a** | Leaves of plants inoculated with the indicated strains. Scale bar = 2 cm. **b** | Nodule phenotype at 21 days post inoculation. Scale bar = 1 mm. **c** | Bacteroid viability determined by live-dead staining of nodule sections and confocal microscopy. Top row images, full nodule sections; Bottom row images, enlarged images of symbiotic cells. Scale bars are indicated in each panel. **d** | Composite image of a *yejF*-infected nodule section showing the rapid permeabilization of the membranes of the internalized bacteria (red staining). Note the bacteria in infection threads (white arrows) are not permeabilized (green staining). The images of the composition are separated by dashed lines.

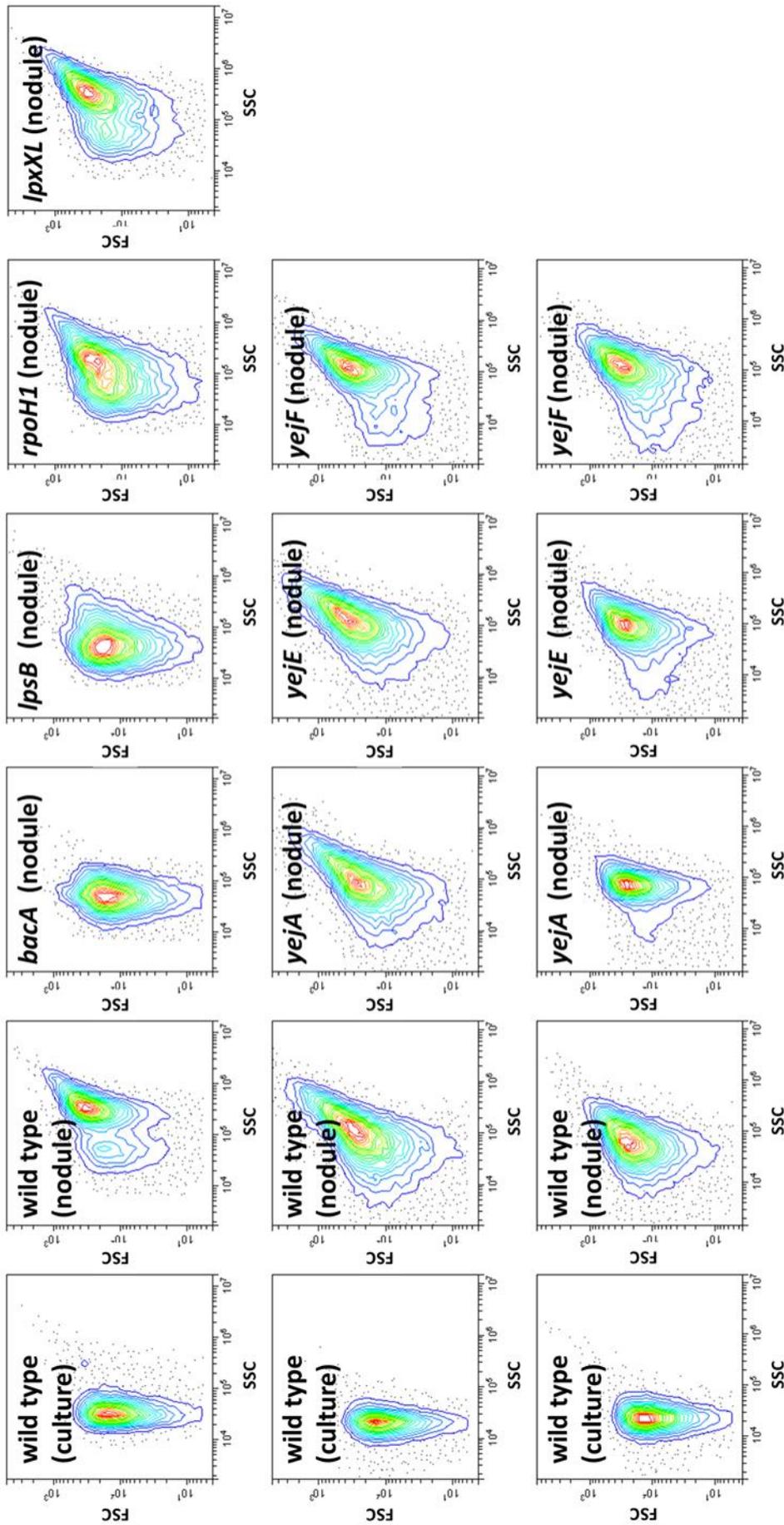


Supp. figure B5 | Morphological parameters in nodule bacteria of *Medicago truncatula* nodules.

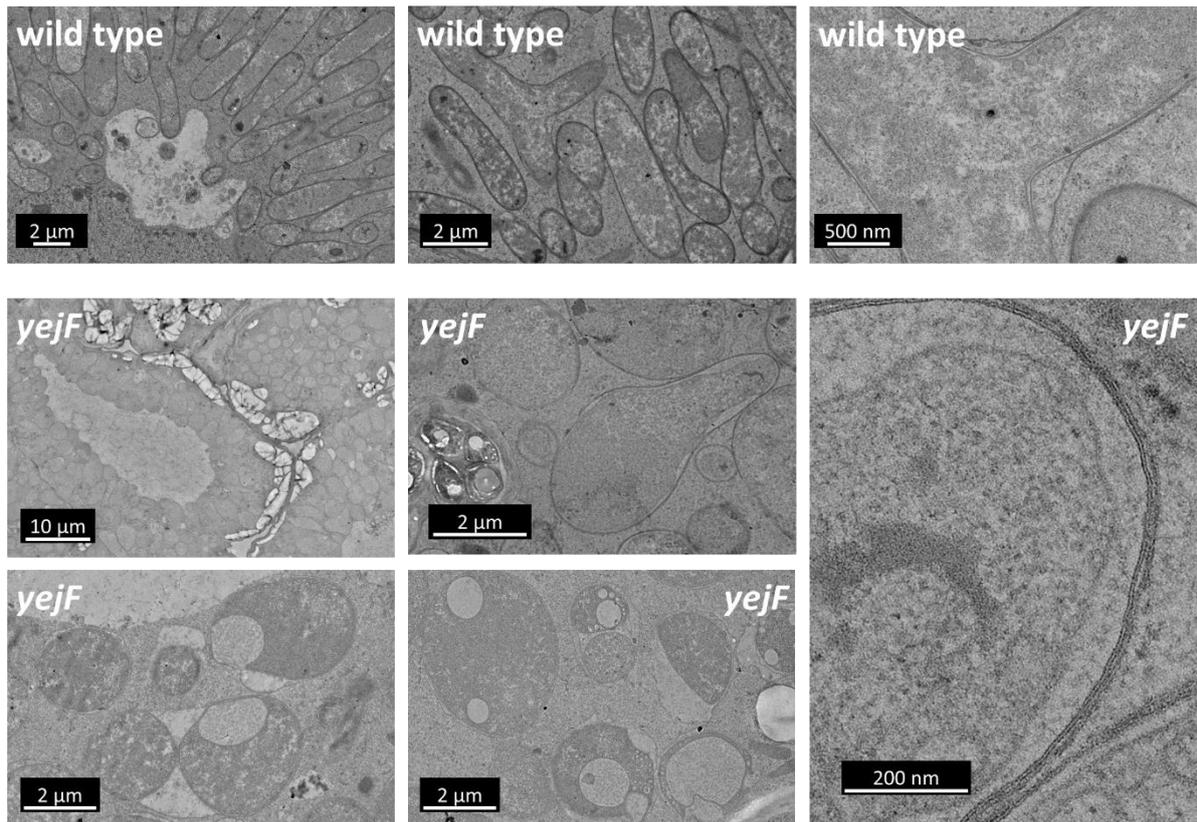
Flow cytometry analysis of the morphology of bacteria isolated from nodules infected with the indicated strains. The forward scatter (FSC, y-axes) is represented in function of the side scatter (SSC, x-axes). The "X" indicates the peak values for each sample. The position of the peak value in the cultured bacteria and in the wild type bacteroids are indicated in each panel by the hatched lines.



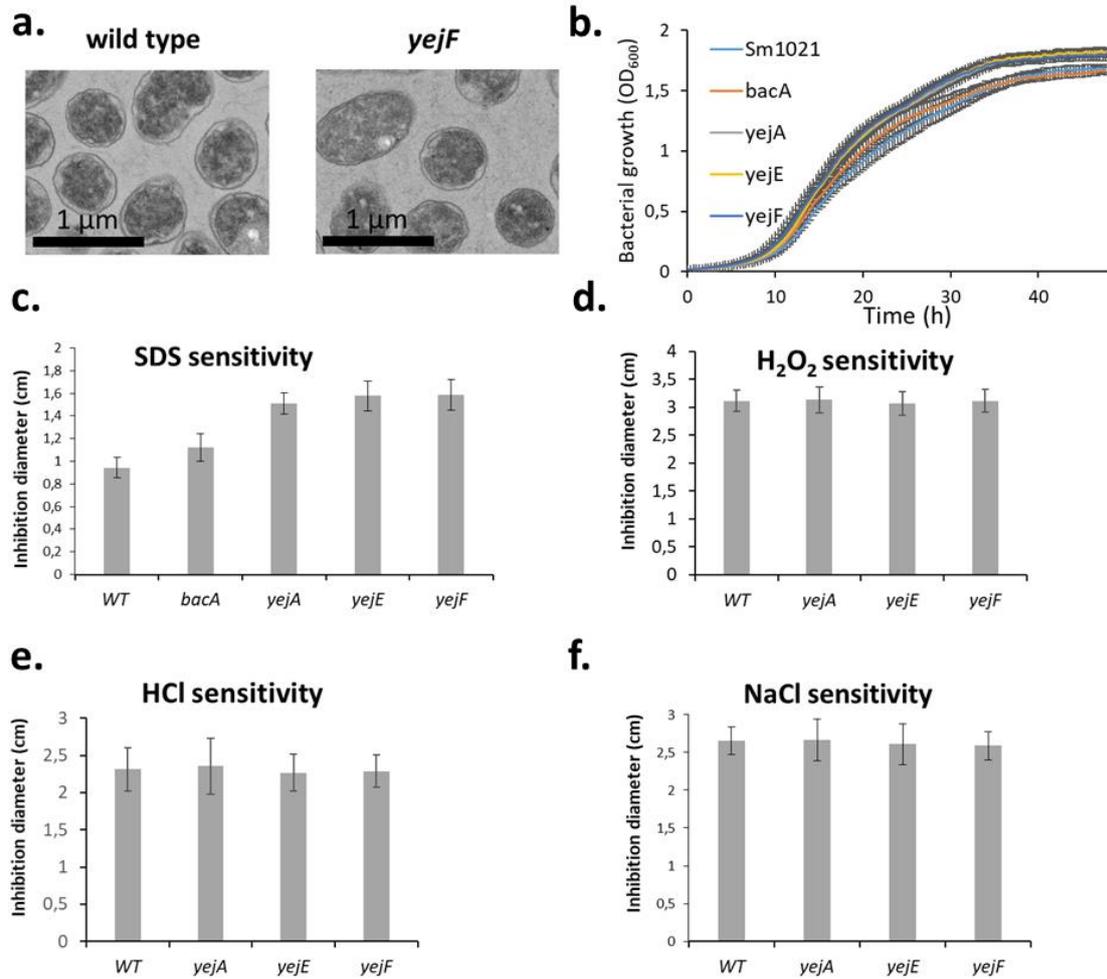
Supp. figure B6 | DNA content in nodule bacteria in *Medicago sativa*. Flow cytometry analysis of the DNA content of bacteria in culture or isolated from nodules infected with the indicated strains and stained with 4',6-diamidino-2-phenylindole (DAPI). The cell counts (y-axes) are represented in function of the DAPI fluorescence (x-axes). The arrow indicates the mean DNA content of wild type bacteroids. The rows represent three sets of independent experiments and the values of DAPI fluorescence cannot be compared between experiments because of differences in instrument settings. The first two rows are samples prepared from 21 dpi nodules; the third row is from 32 dpi nodules. The latter shows that the phenotype of the *yejA* mutant evolves in older nodules.



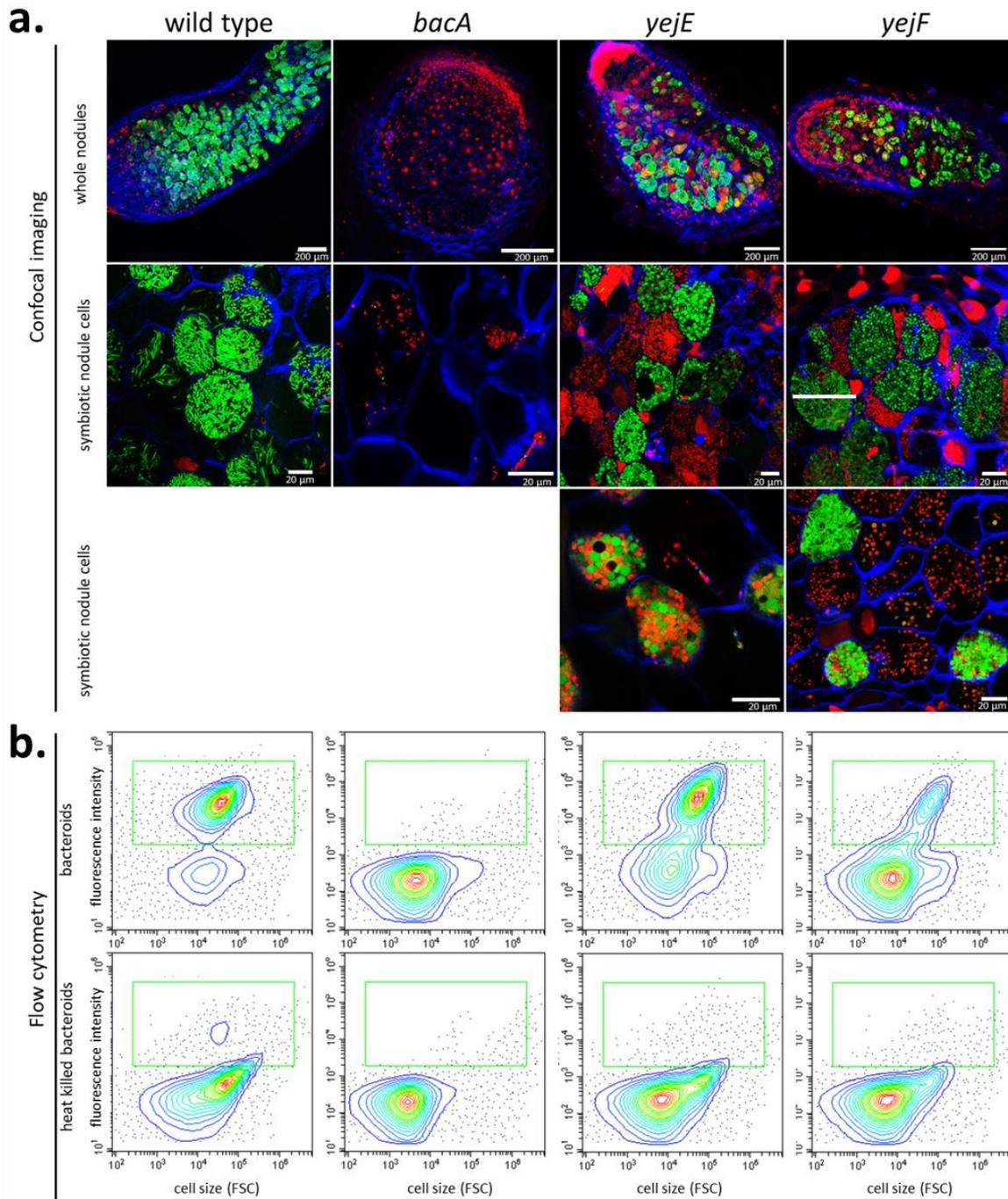
Supp. figure B7 | Morphological parameters in nodule bacteria of *Medicago sativa* nodules. Flow cytometry analysis of the morphology of bacteria isolated from nodules infected with the indicated strains. The forward scatter (FSC, y-axes) is represented in function of the side scatter (SSC, x-axes).



Supp. figure B8 | Membrane defects in the *yejF* mutant bacteroids in *Medicago sativa* nodules.
Transmission electron microscopy of wild type and *yejF* mutant bacteroids.



Supp. figure B9 | Phenotypes of wild-type and *yej* mutant bacteria in culture. **a** | Ultrastructure by transmission electron microscopy. Scale bars are 1 μm . **b** | Growth in YEB liquid medium. **c** | Growth inhibition in the presence of SDS. **d** | Growth inhibition in the presence of H₂O₂. **e** | Growth inhibition in the presence of HCl. **f** | Growth inhibition in the presence of NaCl. Growth inhibitions in **c** to **f** were determined by the size of a halo formed around a round paper patch soaked in the stress compound.



Supp. figure B10 | Nitrogenase expression in the *yejE* and *yejF* mutant bacteroids in *Medicago sativa* nodules. **a | Confocal microscopy of sections of nodules infected with Sm1021.*ppnifH-GFP*, *bacA.ppnifH-GFP*, *yejE.ppnifH-GFP* or *yejF.ppnifH-GFP* and stained with propidium iodide. **b** | Flow cytometry of GFP levels in differentiated nodule bacteria (upper panels) and heat killed nodule bacteria (lower panels). The square shows the position of the GFP-positive bacteroids.**

peptide (concentration)	survival (%) relative to untreated bacteria (n=3)	Standard deviation	survival relative to WT (%)	peptide (concentration)	survival (%) relative to untreated bacteria (n=3)	Standard deviation	survival relative to WT (%)
NCR169 (10µM)				NCR247 (25µM)			
WT	8,72727273	3,40708416	100	WT	0,03025455	0,016354	100
ΔbacA	4,8	4,42920941	55	ΔbacA	0,00116364	0,00272567	3,84615385
ΔyejA	10,9090909	8,51771041	125	ΔyejA	0,04654545	0	153,846154
ΔyejE	9,45454545	1,70354208	108,333333	ΔyejE	0,00116364	0,00272567	3,84615385
ΔyejF	13,6727273	8,85841882	156,666667	ΔyejF	0	0	0
ΔLpxXL	1,01818182	0,34070842	11,6666667	ΔLpxXL	0	0	0
ΔLpsB	6,54545455	1,70354208	75	ΔLpsB	0	0	0
ΔRpoH	7,56363636	0,68141683	86,6666667	ΔRpoH	0	0	0
NCR183 (10µM)				NCR280 (25µM)			
WT	0,04654545	0,02725667	100	WT	0,05352727	0,01090267	100
ΔbacA	0,00581818	0,01362834	12,5	ΔbacA	0,0128	0,02453101	23,9130435
ΔyejA	0,128	0,10902669	275	ΔyejA	0,00116364	0,00272567	2,17391304
ΔyejE	0	0	0	ΔyejE	0	0	0
ΔyejF	0	0	0	ΔyejF	0,00581818	0,01362834	10,8695652
ΔLpxXL	0	0	0	ΔLpxXL	0	0	0
ΔLpsB	0,00232727	0,00545133	5	ΔLpsB	0	0	0
ΔRpoH	0,22690909	0,1499117	487,5	ΔRpoH	0,07214545	0,02180534	134,782609

Table S1 | Assessment of mutants resistance to different NCR peptides. Growth of the bacterial strains analyzed in this study after treatments with NCR169, NCR183, NCR247, NCR280 was assessed by plating and counting.

	Gene	Primer name	Sequence	Description
Insertion mutants	SMc02829 (<i>yejF</i>)	2829forward	CAAATCGGTGTCGACCAACTCGATCCTGAAGCTC	Primers used to amplify a region of <i>SMc02829</i> CDS, flanked with Sall & XbaI sites
		2829reverse	GTTGGCTCTAGACATGGCTATCATTAC	
		2829out-f	CATCTCCTTCGACATCAAG	Primers used for verification
		2829out-r	CTCGTCGGCGATCAGGAGCTC	
	SMc02830 (<i>yejE</i>)	2830forward	CTTCGGCCTGTCGACGTTTGCCGAGTTCATC	Primers used to amplify a region of <i>SMc02830</i> CDS, flanked with Sall & XbaI sites
		2830reverse	CAGCGCTCTAGAGAAAAGGACCGAGATG	
		2830out-f	CTGGCTCTTCTCCTCCTCTTC	Primers used for verification
		2830out-r	GATGATGAGCAGGATGTAGAG	
	SMc02832 (<i>yejA</i>)	2832forward	GCTACGTCGACGATCGGAGAACCCAAATAC	Primers used to amplify a region of <i>SMc02832</i> CDS, flanked with Sall & XbaI sites
		2832reverse	GTGACTCTAGACTCGCTGTCTTTTC	
		2832out-f	CAATGCCGAGGAACAACC	Primers used for verification
		2832out-r	GAGAATGTTCCGGCAGTTC	
SMc00646 (<i>rpoH1</i>)	RpoH1-Sal1-forward	CGCACAGTCGACTAACGAGCCATCTTCGCCTC	Primers used to amplify a region of <i>SMc02829</i> CDS, flanked with Sall & XbaI sites	
	RpoH1-Xba1-Revers	CGAAGATCTAGAGCTCACGATCGTTCAAGAACC		
	RpoH1-up	AAGCGGTATCAGGAG	Primers used for verification	
RpoH1-down	CCTCTGCACCTTTT			
Deletion mutants	SMc02829 (<i>yejF</i>)	2829fup	GCAACGGCACCATCATGGATCCGCACC	Up fragment used for deletion of <i>SMc02829</i>
		2829rup	GGAGAACTCGGAACGTCTTTCTCG	
		2829fdown	CGTCCGAGTTTCTCCCATGC	Down fragment used for deletion of <i>SMc02829</i>
		2829rdown	CACCCACTCGCTCATCCAAGCTTTTCAGG	
	SMc02830 (<i>yejE</i>)	2830fup	GGACACGTGTTCCGCAACGGATCCCTGATC	Up fragment used for deletion of <i>SMc02830</i>
		2830rup	CGTCGCGCCAGACGTCCCTCCG	
		2830fdown	GGACGTCTGGCGGACGATGACAG	Down fragment used for deletion of <i>SMc02830</i>
		2830rdown	GCGAACATGCCTGAAAGCTTCGTCG	
	SMc02832 (<i>yejA</i>)	2832fup	GGTCGAACCGCAGGGATCCAGCGCAGCAGC	Up fragment used for deletion of <i>SMc02832</i>
		2832rup	GCAAGAGCTCCGGCCGTACCCC	
		2832fdown	CGGCCGAGCTCTTGCCTCAGG	Down fragment used for deletion of <i>SMc02832</i>
		2832rdown	CATGATGCCGACGATCGTCAAGCTTATCAGC	

Table S2 | List of primers used in this study.

xiii. Additional results and discussion

These results will not be part of the final submitted manuscript. Procedures used to carry out these complementary experiments are identical to those presented in the paper.

Roles of *yejE* and *yejF* genes in NCR resistance in planta

In order to better understand the role of *S. meliloti* Yej transporter, we have used the *dnf1-1*, *dnf4* and *dnf7* *M. truncatula* mutants as tools to dissect the development of nodules elicited by the mutants. Indeed, each of these three mutants have altered NCR content. The *dnf1-1* mutant cannot secrete any NCR peptide and *dnf4* and *dnf7* respectively lack NCR211 and NCR169. As previously shown in **figure 23**, NCR211 is expressed in the intermediate zone (II-III / IZ) whereas NCR169 is detected both in intermediate and in fixing zones (IZ / ZIII). From these expression patterns, we could expect NCR211 is important during the differentiation while NCR169 could play part both during differentiation and nitrogen fixation steps. Using both mutants with the *dnf1-1* lineage, we hoped to find out the stage of TBD in which *yej* transporter is required (**Fig. 42**).

As it has been previously shown, bacterial *yej* and *bacA* mutants survive in *M. truncatula* *dnf1-1* mutants with no apparent alteration of the morphology. Although *dnf4* and *dnf7* plant mutants developed nodules much slower than WT plants, they could become mature and induce TBD of *S. meliloti* WT bacteroids in our hands (**Fig. 42**). Cell death was observed in the fixing zone, though in a lesser extent than previously reported (Horvath et al., 2015; M. Kim et al., 2015). The *bacA* mutant survives in *dnf7*, even if it still rarely and badly colonizes plant cells. In *dnf4*, half of *bacA* bacteroids seem to die in the few infected cells. As expected, in both mutant plants, $\Delta bacA$ does not undergo TBD. Though $\Delta bacA$ seem to better survive in *dnf4* and *dnf7* mutants, they seem quite affected even with the lack of either in NCR211 or NCR169. As described in the literature and in this study, the BacA transporter is at least involved in the transport of NCR247 (Guefrachi et al., 2015) and might also be able to transport other NCR peptides. Therefore, if BacA could play part in the resistance to NCR211 and NCR169, the death and colonization problems in *dnf4* and *dnf7* only reflects the importance of BacA for the transport or interaction with other peptides as well.

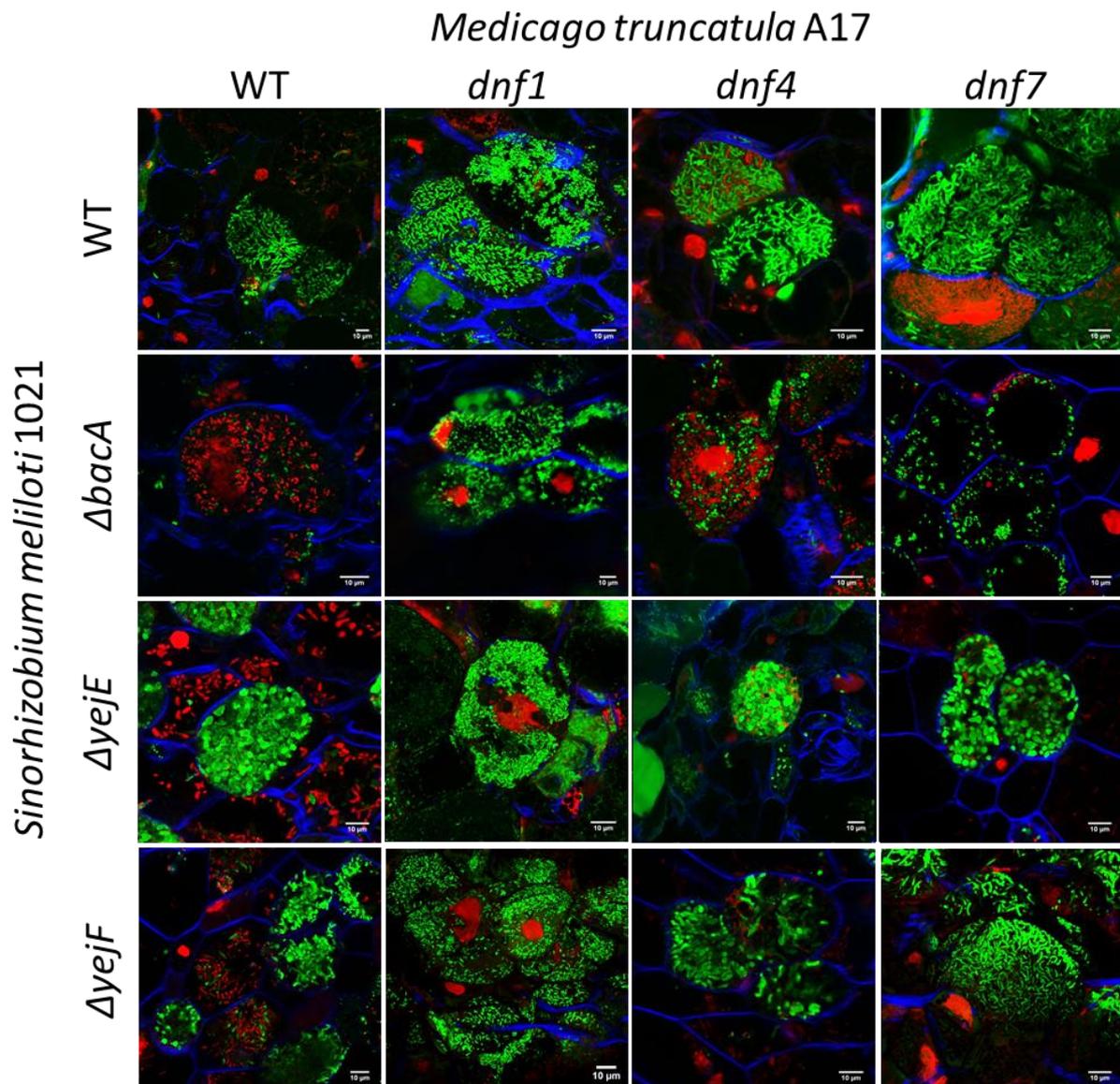


Figure 42 | Symbiotic phenotype of *Sinorhizobium meliloti* 1021 mutants during symbiosis with *Medicago truncatula* *dnf* mutants. Bacteroid viability assessed by live-dead staining of nodule sections and confocal microscopy. Various combination of plant (columns; WT, *dnf1*, *dnf4* and *dnf7*) and bacteria (rows; WT, *bacA*, *yejE*, *yejF*) mutant couples were assessed. Enlarged images of symbiotic cells are shown. Scale bars are indicated in each panel.

The *yejE* mutant displays partially altered morphology both in *dnf7* and *dnf4* mutant plants with swollen bacteroids, which are thicker than their WT counterparts. The *yejF* mutant morphology appears to be similar to *S. meliloti* WT in *dnf7* whereas in the *dnf4* mutant its bacteroids become thicker (**Fig. 42**). In any case, when the abnormal morphology is observed, it does not imply cell death as in *M. truncatula*. As *yejF* bacterial mutant loses its abnormal morphology in *dnf7* mutants, it is possible *yejF* is involved in the resistance or perception of NCR169. As this behavior is not observed with *yejE*, we can assume the *yejB* permease is

enough to maintain a functional permease domain for NCR169 transport. Perhaps various types of *yej* transporters coexist in bacterial envelopes, displaying variations in the transmembrane domain organization: YejB homodimers, YejE homodimers and YejBE heterodimers. Each of these importers could have variable affinities with NCR peptides. Structure of this type of transmembrane proteins in complex is difficult to resolve due to their hydrophobicity. Co-immunoprecipitations would encounter the same issue. However, co-localization or two-hybrid assays using translational fusions could help resolve this question.

The fact that $\Delta yejF$ seems to differentiate normally in *dnf7* nodules suggest this mutant is still able to interact with the NCR peptides and to initiate TBD, but can resist the lysis and overgrowth that happens in *M. truncatula* WT. This feature is interesting as it suggests certain NCR only or a certain dose of NCR could trigger the *yej* phenotype. In a similar way that cationic NCR might permeabilize bacteroids so a second wave of NCR peptides can trigger the TBD, it could be possible that a third wave exist to provide negative regulation on the TBD. This function might help stopping the differentiation process in order to maintain bacteroid integrity. To do so, NCR from this hypothetical third wave could interact with TBD-inducing and permeabilizing NCR to prevent their action. Indeed, it has been shown that NCR peptides can physically interact with each other in bacteroids (Farkas et al., 2014). An alternative or complementary hypothesis is that the *yej* transporter could be involved in the maintenance of membrane integrity in the nodule. This assumption would be coherent with the increased SDS sensitivity observed in the *yejF* mutant. However, this hypothesis alone does not explain the inability of $\Delta yejF$ to transport NCR247.

YejABEF and BacA transporters have been addressed for their essentiality in symbiosis. Are they specifically dedicated to managing NCR during symbiosis or do they play part in life in the soil? They seem to be induced in nodules and more specifically in zones I and II for *yejABEF* and from zones I to II-III for *bacA* (Roux et al., 2014), but their expression levels in soil was not assessed. Still, they could be important for managing other types of threats such as bacterial AMP or toxins. Functional study of a microbial community could be carried out though needing to reconstitute a *M. truncatula* core soil microbiome.

yejABEF genes seem to be conserved in *Bradyrhizobium* sp. ORS285 and *B. diazoefficiens* USDA110 as shown by synteny analysis (**Fig. 43** ; <https://mage.genoscope.cns.fr/>

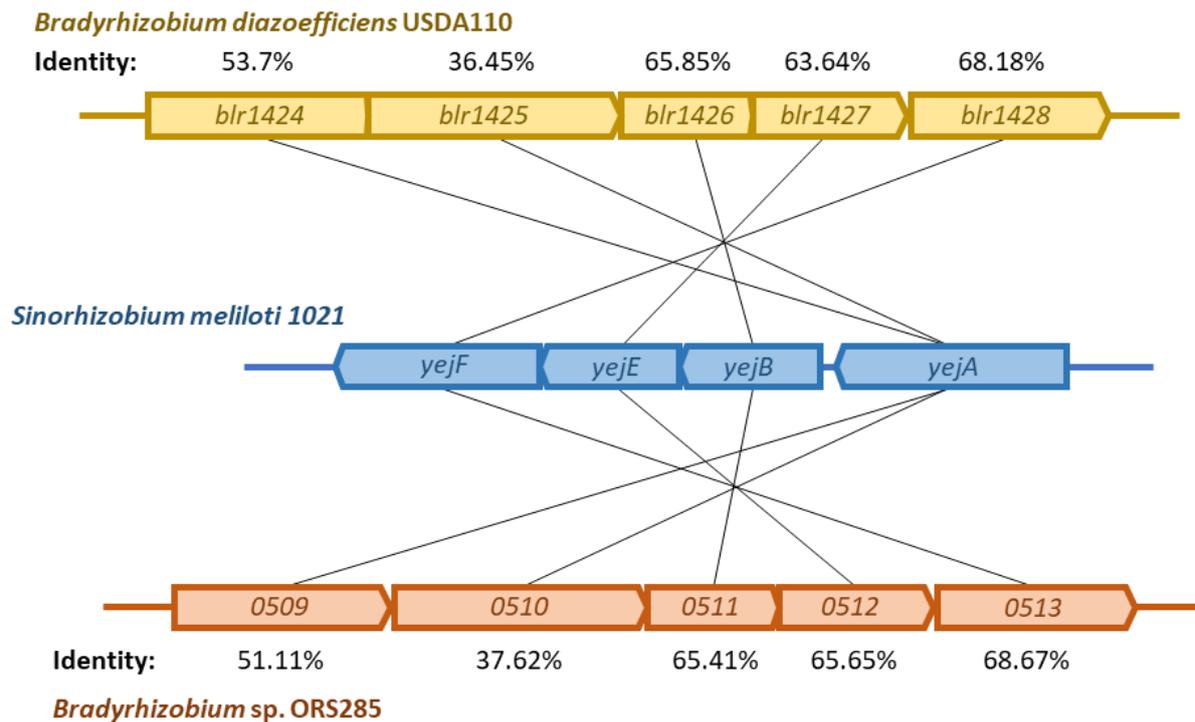


Figure 43 | Conservation of *yejABEF* genes synteny. Gene order of the *yejABEF* genes is conserved in other rhizobia studied in this thesis. The identity percentage calculated from the amino-acid sequence is specified for each ortholog. Gene id of ORS285 orthologs was simplified for graphic representation reasons. Their full id can be obtained by adding "BRAD285_v2_" before each number.

- MaGe). It would be interesting to investigate their role in both bacteria for the resistance to NCR peptides. As shown in Barrière and colleagues (2017) loss of *bcIA* in USDA110 does not lead to symbiotic defect even in a Dalbergioid plant. Though it is thought USDA110 envelope lipid composition is largely responsible for its resistance to NCR peptides, whether the *yej* transporter would contribute to it cannot be excluded.

How do BacA and YejABEF importers carry out their activities?

BacA and YejABEF transporters are supposed to be both located on the inner membrane as ABC transporters usually are in Gram-negative species. This means NCR must pass through the OM in order to be imported. However, it is unsure how NCR transit through the OM. Usually, Gram-negative bacteria harbor pores and non-selective channels on their OM which allow transit of various molecules in the periplasm (Nikaido, 2003). Some molecular systems have been described in Gram-negative bacteria to be able to transport molecules from cytosol to extracellular environment with a single multimeric complex. This is what the type one secretion systems do, using the TolC OM protein (Davidson et al., 2008). Interestingly,

though being usually associated with export activities, TolC was also shown to import peptides such as Colistin (Zakharov et al., 2012). One can wonder if such system is used to manage the NCR stress around the envelope. Interestingly, a *tolC* mutant in *S. meliloti* is unable to perform functional symbiosis by forming white nodules which seems to be badly colonized (Cosme et al., 2008). Though the symbiotic phenotype of the *tolC* mutant is restrained to this analysis, it resembles the $\Delta bacA$ phenotype. It could be explained by an interaction between BacA and TolC to transport NCR peptides. As the phenotype of the *tolC* mutant is much stronger and different from *yej* phenotype we can assume *yej* function is different from that of *tolC*. Moreover, *yejA*, which is annotated as a PBP and confirmed as such by structure prediction (Phyre² - data not shown), seems to play part in NCR resistance. Therefore, even if TolC could play a role in NCR import by interacting or not with BacA, NCR peptides might have another mean to enter the periplasm. It is known that some mammalian and insect defensins can form large pores in OM and IM which lead to the depolarization of IM, cessation of respiration and lysis (Letellier & Bonhivers, 1996). NCR peptides have membrane permeabilization activity, though their exact mechanism of action is unknown. This could be their way in the periplasm. This hypothesis is not incompatible with the previous one related to *tolC*, as interaction of this protein with BacA does not explain alone the role of *yej* transporter in NCR resistance. Moreover, given the large array and doses of NCR peptides bacteroid must face, it is likely that, even if the described transport mechanisms could reduce the NCR charge on the envelope, they might not be enough to nullify it, though they could help reduce it to non-damaging levels.

Specificity of NCR action

Though this analysis is restrained to two single NCR-lacking mutants, it might have given clues on the roles of *yejABEF* transporter during symbiosis. It is a chance that such features were observed with NCR169, given the diversity of NCR, and the fact they probably act jointly or are strain-specific. So far, two NCR peptides were described to be having strain-specific actions. They are the NFS1 and NFS2 peptides, which alleles can prevent the colonization by specific bacteria (Gourion & Alunni, 2018; Q. Wang et al., 2017; S. Yang et al., 2017). In this case, it seems they are used by the plant to restrain its host range. This mechanism could help the plant avoiding cheaters, pathogens, or even just inefficient nitrogen fixers. Given

the diversity of bacteria might face in the soil, it is possible they have developed arrays of peptide which allows to prevent colonization by improper bacteria. However, it seems legume plants possess other way to control the entry of strains in their tissues. A plant mutant in the *apn1* displays a proper colonization though associated with defense reaction and early senescence of *Lotus japonicus* nodules elicited by *Mesorhizobium loti* (Yamaya-Ito et al., 2018). This phenotype is similar to those of the *NFS1*⁻ and *NFS2*⁻. The mechanics behind this gene functioning are unknown, however, it encodes aspartic peptidase which is possibly targeted to the symbiosomes. Though aspartic peptidases were shown to have antimicrobial properties when carrying a Plant Specific Insert domain (Cheung et al., 2020), *apn1* do not harbor such domains. It is thus closer to aspartic peptidases such as *CDR1* in *Arabidopsis thaliana* (Xia et al., 2004) which is thought to be involved in disease resistance signaling. It seems *M. truncatula* possess an ortholog of *apn1*, though not characterized yet (Gourion & Alunni, 2018). The coexistence of two mechanisms in IRLC to control host range is thus possible, one relying on NCR peptides and the other on defense reaction signaling.

GENERAL DISCUSSION AND CONCLUSION

i. Cell envelope integrity – a key feature for a successful symbiosis

As described in our study of bacterial functions and their role in NCR resistance, the envelope seems to harbor multiple facets that would play part in the TBD process. Other functions described in the introduction (hopanoids, EPS, peptidoglycan) are important for enduring NCR peptides. The ability of USDA110 to survive within an NCR-producing plant whereas it is supposedly not adapted to such stress is likely due to its envelope composition. Indeed, up to 40% of USDA110 lipid fraction is composed of hopanoids (Kannenberg et al., 1996), therefore protecting USDA110 from any lethal membrane damage. Apart from these functions, membrane microdomains could play part in envelope structuration and therefore in resistance to NCR peptides. Different types of membrane microdomains have been described in bacteria. Cardiolipin is a specific subtype of lipids composed of two phosphatidyl moieties linked together by a glycerol, their structure is such that they induce bending of a membrane. They typically are responsible for the convolutions of mitochondria inner membrane or morphology of rod shaped bacteria and notably in rhizobia (Mileykovskaya & Dowhan, 2009). Interestingly, cardiolipin can form membrane microdomains. As *S. meliloti* produces cardiolipins (Vences-Guzmán et al., 2008), an alternative hypothesis to *yejABEF* mutants altered morphology is these domains are disordered. Localizing cardiolipin domains thanks to specific fluorescent lipidic probes would tell whether these domains are intact (Leung et al., 2013). Combining this experiment with a localization of YejABEF would help see whether the transporter is microdomain-specific or not. Moreover, we can wonder how the cardiolipin content and microdomains are affected in spherical bacteroids. Reorganization of such domains might be a key step in the transition between elongation and spherization of such bacteroids.

Another type of membrane microdomains are lipid rafts (LR). These domains are pretty well characterized in eukaryotes but related knowledge in bacterial biology is scarce. Though having been considered as experimental artifacts for many years, multiple experimental evidence now point out their importance in organizing membrane functions as well as biotic interactions in bacteria (Lopez & Koch, 2017). Being formed of unsaturated fatty acid, sterol-like molecules and lipid-A, LR locally alter rigidity and fluidity of the membrane. These membrane microdomains do not rely only on lipids to perform their functions, but also on

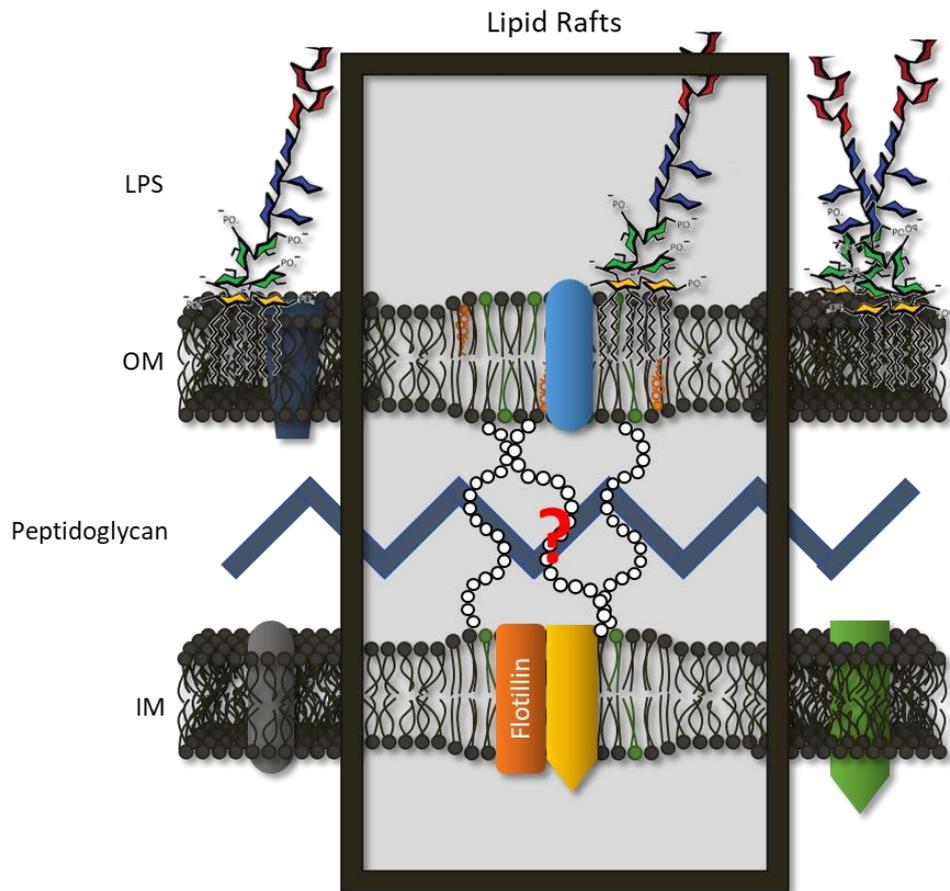


Figure 44 | Suggestion of lipid raft (LR) organization. On the IM, LR could assemble thanks to the presence of flotillins and of unsaturated fatty acids (green lipids). On the OM, hopanoids (orange sterol-shaped molecules), LPS with or without hopanoids could cluster to form LR. Peptides or membrane fusion proteins (white polymers) could help synchronizing function on IM and OM LR of bacteria both locally and temporally.

proteins called flotillin (Lopez & Koch, 2017). The role of these proteins is to recruit other protein in lipid rafts, clustering enzymes involved in a same signaling or synthesis pathway. They also seem to be involved in LR dynamics such as rearrangement of LR across the membrane (Lopez & Koch, 2017). In bacteria, it is possible that OM LR are constituted in majority by hopanoids and LPS, whereas IM lipid rafts would be made of hopanoids and flotillin. Though it has not been described, these proteins could also be a way to organize functions requiring coordination of both IM and OM such as secretion or influx/efflux activities (**Fig. 43**).

It has been shown that both cholesterol and hopanoids can self-assemble into LR (Sáenz et al., 2012). Hopanoids were shown to be essential in symbioses involving *Bradyrhizobium* species (Kulkarni et al., 2015; Silipo et al., 2014). Indeed, hopanoids rigidify membranes, which help resisting various envelope stresses *in vitro*. Besides the presence of NCR peptides, it is

known that the nodule environment can be harsh and therefore hopanoids are essential for the establishment of a successful symbiosis. Hopanoid might prevent permeabilization of the bacterial envelope by NCR peptides. The role of 'free' hopanoids was however impossible to dissociate from that of the LPS-linked ones. Identifying the gene responsible for the linkage of hopanoids to the lipid A could be a critical step in the search of LR and analysis of their composition. Indeed, though in some bacterium both hopanoids and LPS seem to cluster in LR, the presence of LPS-linked hopanoids could mean these molecules and standard LPS distribute differentially into LR and the rest of the membrane.

Though not being able to synthesize hopanoids, *S. meliloti* may possess an ortholog of these flotillin (SMb20989; 21% and 24% identity with *Bacillus subtilis* or *Methylobacterium extorquens*). Do LR exist in *S. meliloti* and do they play part in symbiosis? This is yet to be discovered but techniques exist to or to visualize LRs or to extract detergent resistant membranes (DRM). DRM represent what can be extracted through biochemistry (including LR and artifacts). During my Master2 internship I have worked on the detection and characterization of LR in *Bradyrhizobium* sp. ORS285 and successfully isolated DRM. LR are usually found in these fractions that are usually extracted from eukaryotic cells. Radiometric imaging to visualize their presence over the cell was also used to try corroborating these observations. Many other arguments are needed to prove the existence of lipid raft in ORS285 though. Flotillin-like localization and role(s) are one example. Do NCR peptides first require docking to LR thanks to flotillins or LR lipids? Investigating the existence and the role of membrane microdomains in rhizobia may give interesting insights on the mechanism of action of NCR peptides and their related transporters in symbiosis and TBD.

ii. Roles of NCR peptides and evolution of TBD

NCR247 treatment induces cell elongation *in vitro* though not as strong as *in planta*. It is likely that treating a bacterial culture with a cocktail of NCR induces a much stronger response as one can imagine different NCR peptides might interact with various targets in the cell-cycle regulation, as well as other envelope and metabolic functions. More than 600 *NCR* genes are found in *M. truncatula* often cluster into specific genomic locations. It is supposed both gene duplication and exchanges of gene pieces between paralogs happened in these regions which would result in a dynamic expansion of the *NCR* gene repertoire (Alunni et al., 2007; Kondorosi

et al., 2013; Pecrix et al., 2018). The high diversification these peptides undergo and the convergent evolution leading to the acquisition of these genes independently in Dalbergioid for example could point out the essentiality of their action. Though it seems TBD yields an increased symbiotic efficiency, one can wonder why such functions were acquired. NCR peptides could be acting as a filter used to discriminate between efficient and inefficient strains as suggested by the study of NFS1 and NFS2 NCR peptides (Wang et al., 2017; Yang et al., 2017). Was it for the plant to increase nitrogen fixation yields from formerly inefficient bacteria or to avoid pathogenic and/or saprophytic colonization of the nodule? Or else was it that thousands of years ago rhizobia were so rare that after acquiring one symbiont plants needed to get the maximum from it? Indeed, from primitive agriculture to modern agriculture we might have deeply altered rhizobia geographical repartition through trade, colonization, and importation of cultures in unnatural habitats without even noticing. Moreover, it has been shown that recently generated soybean cultivars are less efficient in discriminating non-symbiotic bacteria, or cheaters from its symbionts than ancestral cultivars (Kiers et al., 2007). In depth analysis of both rhizobial and legume plants phylogeography could help answering those questions.

In the end one can wonder if this acquisition is going to persist or decline. Is the acquisition of NCR a first step in a transition to obligatory symbiosis? This seems unlikely as building an obligatory endosymbiosis requires transfer to the offspring and transporting rhizobia from roots to seeds which seems challenging. In aphid – *Buchnera* symbiosis, the vertical transmission is possible as the symbiotic organ is located near the embryos, which allows a migration of the symbionts for the infection the developing embryo as early as the gastrula stages (Koga et al., 2012). Nevertheless, microbial communities can be found in every organ of plants (Bulgarelli et al., 2013), flowers included. In 2014, Mora and colleagues succeeded in isolating nitrogen-fixing rhizobia from *Phaseolus vulgaris* seeds. The bacterium, a *Rhizobium phaseoli* strain, has a peculiar metabolism though, resulting from its lifestyle inside the seeds. Even though being able to elicit nodules on bean and fix nitrogen, no study demonstrates whether this could happen naturally. In the lab, we possess a batch of alfalfa "*Bella Campagnola*" seeds which are infected by WT rhizobia. It seems possible a "vertical transmission" happens naturally, but it is unsure if is a widespread phenomenon or specific to certain growth conditions or host/microbial genotypes. This would require a colonization of

the embryo surface, which leads to the presence of rhizobia close to the shoot apical meristem that will disperse rhizobia on the shoot. Rhizobia therefore would need to be able to maintain on such environment. Considering this to be possible, bacteria would require colonizing the floral meristems in order to infect future embryos. Each of these key step makes it difficult to imagine a transition to an obligatory and vertically transmitted symbiosis. Moreover, genome reduction is usually a marker of dependency of the symbiont for its host, such as in aphids – *Buchnera* symbioses (Koga et al., 2012). In the above presented scenario, genome reduction is unlikely to occur as the bacteria must adapt to various environments. It therefore unlikely that such symbiosis might become obligatory in short or medium term (with an evolutive scale). Several evolutive steps are required but perhaps in hundredth of thousands of years this symbiosis will have evolved vertical transmission features.

iii. AMPs control biotic interactions

The above-discussed evolutive fate of NCR peptides is unclear today but some homologous AMPs have been described, having similar roles in managing symbiotic interactions, and sometimes in more intimate ones. In stinkbug – bacteria symbioses a family of peptides which strikingly resemble NCR was identified (Futahashi et al., 2013). These peptides called CCR (Crypt-specific Cysteine-Rich) peptides similarly to NCR peptides, resemble defensins and are specifically expressed in the symbiotic organ (Futahashi et al., 2013). Though their function is yet to be fully understood, it is supposed they control the entrance of bacteria into the symbiotic organ and/or control multiplication within the organ. Other plant symbiotic interactions have been shown to display similarities with the rhizobium–legume symbiosis. Though not extensively described in the manuscript, the formation of nitrogen-fixing nodules is not restrained to legumes. Actinorhizal plants such as *Alnus*, can be infected by Gram-positive bacteria of the *Frankia* genus. Though the roles are not fully explored yet, remarkably *Alnus* can secrete AMPs, which seem to permeabilize the bacteria in order to facilitate metabolite exchanges (Carro et al., 2015).

Subsequently to the permeabilization of *Frankia* membrane, the *Alnus* AMPs have been shown to induce the release of different metabolites from the cells, especially specific amino acids that may contribute to the plant nitrogen nutrition. Similarly, in aphid – *Buchnera* symbiosis, AMP peptides called BCR (Bacteriocyte-specific Cysteine-Rich) are secreted by the

host, also having membrane permeabilization activities. Possibly due to genome reduction of the symbiont, few transporters are produced to perform the metabolic exchanges between host and symbiont (Mergaert et al., 2017). Therefore, instead of transporters, BCR might be performing this activity. As shown on **figure 17a**, several transporters could facilitate the dynamic flux of metabolites between partners in the interaction. As an example, *S. meliloti* possesses homologs of such genes, which seem to be induced *in planta* (Roux et al., 2014). Contrarily to aphid – *Buchnera*, NCR peptides might not be the only ones to be responsible for metabolic exchanges, but they may still facilitate them.

iv. The symbiotic genes quest

As described in the result chapters, finding good candidate genes which function is required during symbiosis is a first critical step in discovering efficient symbiotic systems. Two approaches can be used to assess new functions: with or without *a priori*. Without *a priori* approaches are required to unveil functions that are not described in related biological systems or to overcome a subjective search for new functions. In this kind of approach, predefined criteria are used to choose targets. Though transcriptomic analyses are great to dissect the regulation network of an organism, a tissue, or a cell, they show their limits when used to

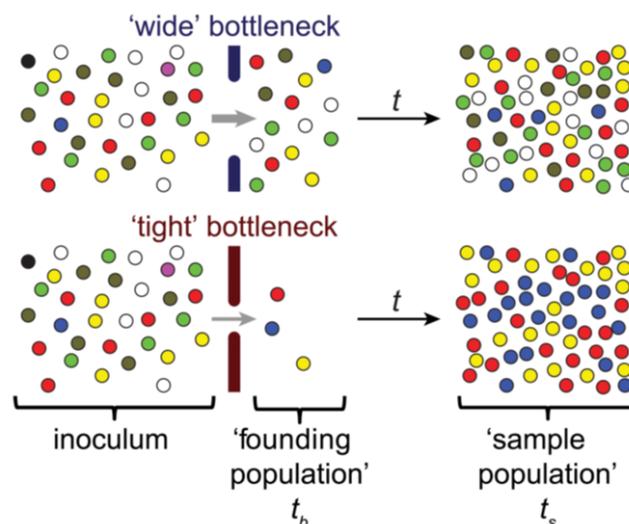


Figure 45 | Schematic representation of bottleneck effects. Spheres represent individual bacteria and their color display genotypic diversity. The schematic gates drawn as solid bars represent either wide or tight bottleneck depending on the size of the gap. A bottleneck will reduce the size of the infecting population. The resulting founder population will replicate, giving rise to the sample population. As described on this scheme, wide bottlenecks do not alter the genetic composition of the population, or only in a limited extent. The sample population is similar to the inoculum in this situation. In contrast, a tight bottleneck severely alters the genetic makeup of the sample population because of a stochastic sampling of the inoculum. Extracted from Abel et al., 2015.

predict gene essentiality. Similarly, proteomics shows similar traits as being based on the difference of proteins accumulation to identify targets. Moreover, proteomics does not give a full picture of the bacterial physiology, as protein extraction protocols specifically enrich for specific protein types (usually cytosolic proteins, passing by most membrane proteins). Metabolomics are interesting to overview the metabolic capacities of system though being disconnected from signaling. In addition, when analyzing the content of metabolites in nodules, it is difficult to know from which partner they come. Transposon insertion sequencing (TIS) techniques are a good alternative as they directly indicate whether a gene is essential in the test condition.

However, this approach also has some limitations, and its main issue is its high sensitivity to bottlenecks. A biological process or system, which leads to a strong constriction of the population size, is called a biological bottleneck. Such reduction in a population size can cause stochastic sampling of certain genotypes in the population which disconnect the frequency of a mutant in the library from their fitness (Abel et al., 2015). Sadly, test conditions which are often subject to severe bottleneck are host microbe interactions such as pathogens and symbionts. In these interactions, starting from small founding populations, microbe often grow enormously within their hosts (Abel et al., 2015; Chao et al., 2016). Particularly, during the infection of legume plants by *Rhizobium* bacteria the populations passes through a sharp bottleneck. It is thought nodules are infected by nearly clonal population of rhizobia (Goormachtig et al., 2004; Kondorosi et al., 2013; Remigi et al., 2016).

The simplest way to estimate these bottleneck sizes is through counting the number of initial and harvested cells after host infection (Abel et al., 2015). However, a genetic composition of the population can be altered after having passed through the bottleneck and this technique does not allow to estimate this alteration. Using Tn-seq, this parameter can be taken into account if an inoculum with precise density is used and number of unique insertions measured after host colonization (Brooks et al., 2014; Stephens et al., 2015). This technique could help estimate how many nodules should be harvested to minimize the bottleneck effects.

Due to these bottlenecks, few examples of TIS related to legume-rhizobia symbioses are found in literature (Fabian et al., 2020). Most of them were based on *in vitro* assays. Notably, the study of Arnolds and colleagues (2017) assessed the resistance of *S. meliloti* for the

resistance to NCR247. Very recently, a bacterial gene essentiality for the symbiotic process was assessed through INSeq techniques (Wheatley et al., 2020). They compared the sets of essential genes of *Rhizobium leguminosarum* in rhizosphere, root, and nodule colonization of pea plants. They harvested around 150.000 nodules to be able to conduct the experiment. Over 15 target genes identified as essential in the INSeq experiment, only one was not confirmed after various functional analyses. Despite being quite heavy, this technique seems to yield good results on the identification of genes of functional importance.

v. Assessing symbiotic efficiency

The applied goal of such studies is to find/design strains or couples of bacteria and plants for which legume fruits yields is increased. Symbiotic efficiency is often assessed as the overall plant return on investment which can be measured in its simplest way as the gain in biomass. Consequently, increased efficiency could be associated to increased yields due to energy saving, and decreased nitrogen limitation. As presented in the introduction various parameters ranging from biomass accumulation or nitrogen fixation are used to assess symbiotic efficiency. One parameter of interest which is rarely measured in the literature is seed yield. However, it has been shown in soybean that biomass is positively correlated with seed yield (Rotundo et al., 2014). Taken together these parameters reflect the symbiotic efficiency from an agronomical point of view.

This concept can also be seen from an ecological point of view. Indeed, mutualistic symbiosis concerns interaction where both partners benefit from their association. There usually a dynamic balance between gains and losses both partners can undergo. Indeed, through evolution, the acquisition of genotypes favorizing one in the detriment of the other can be expected. The previous ways to look at legume-rhizobia was host plant-orientated but what do symbionts gain from this interaction? Indeed, TBD through its enslavement of bacteria in profit of the plant can question the advantage of bacteria in this interaction. Actually, they benefit from a controlled environment with nutrient supplies. Even when there is TBD, some undifferentiated bacteria persist in infection threads as well as in invasion and intermediate zone of indeterminate nodules. These can multiply and colonize the soil when the nodule is senescent. Moreover, as previously discussed, some rhizobia can take advantage of the senescence process by invading the saprophytic zone where that can multiply by feeding on

the lysed nodule tissues (Timmers et al., 2000). In the end, these bacteria might end up in the soil again and in higher numbers than if they had to grow in the soil. Therefore, persistence of bacteria in the soil could be an indicator of a functional symbiosis. It has been shown that nodulation in a crop usually leads to an increase of the symbiotic rhizobia in the soil (Thies et al., 1995).

However, these assertions are true considering indeterminate nodules in which TBD occur. In the case of determinate nodules of NCR-like producing plants such as *Aeschynomene* species, whether gain balance for both partners is in equilibrium is less clear. Indeed, we do not know if bacteria can be released during nodule senescence as no infection zone is maintained in this type of nodule. *Aeschynomene afraspera* is able to develop outgrowth on nodules elicited by *Bradyrhizobium* sp. ORS285 in which viable undifferentiated bacteria are hosted but it is one known exception among all *Aeschynomene* species (Bonaldi et al., 2011). In this case, why was this type of symbiosis selected if it is an evolutive dead end? Either we miss an advantage for the bacterium, or this interaction is under an evolutive transition. There might be some features yet to be unveiled in order to answer this question.

Anyway, as in legume-rhizobia symbionts are transmitted vertically, we could extend the concept of symbiotic efficiency to the availability of a bacterium in the environment of the legume plant. This feature could be dependent of various parameters. If we consider symbionts are already present in the soil, initiating a symbiosis means soil and then root colonization by bacteria, as well as competition with other microorganisms. The ability of having efficient chemotaxis to follow the flavonoid trail may also plays part in an efficient colonization. Whereas protocols are available in the literature for competition assays and soil colonization using serologic, molecular, or cytological techniques, assessing root colonization could be more difficult. A way to measure it would be to inoculate in vitro plantlets with fixed amount of bacterial inoculum and measuring after a given time the population found on the root, nodules excluded.

vi. Increasing symbiotic efficiency

It is apparent that multiple factors modulate symbiotic efficiency. As shown by Oono & Denison in 2010 but also by Sen & Weaver in 1984, TBD leads to increased nitrogen fixation

which can be assimilated to a gain of symbiotic efficiency. Several traits in the bacterium and in the plant could be engineered to increase the symbiotic efficiency. One of them is bacteroid morphology (Lamouche et al., 2019a; Lamouche et al., 2019b). Spherical bacteroids seem more efficient than their elongated homologs in symbioses involving *Aeschynomene* species. This is interesting as surface-volume ratios are usually greater on a capsule shape than on a sphere, meaning exchange should be facilitated. However, as shown by our measures of ORS285 bacteroid length in *A. afraspera*, ORS285 reaches a mean length of 4 μm and a mean width of 1 μm . In these conditions, an estimated calculation of its volume and area respectively yield 10.47 μm^3 and 25.13 μm^2 . In comparison a spherical bacteroid of 2.5 μm would have a volume of 65.44 μm^3 and a surface of 78.53 μm^2 . Though with these values, the surface-volume ratio is twice greater in elongated bacteroids than in spherical (2.4 and 1.2 respectively) the much higher volume of the cell seems to give indisputable advantage.

Another way to increase bacteroid efficiency is to alter the regulation of the cell cycle machinery. In the lab, current studies of bacterial cell cycle mutants display a quicker TBD and plant biomass increase compared to the WT strain. However, the control strain eventually seems to catch up these mutants in terms of plant biomass (unpublished data). Perhaps that trying to push symbiotic efficiency further by engineering bacteria only reaches its limits because of different types of constraints. There could be bacteria-dependent limitations, for instance lethality such as in the cell-cycle mutants. Indeed, altering the cell cycle has major impacts on the bacterial life and a few cell-cycle mutants are lethal. It has been observed in the lab (unpublished data) that in plasmid-based over-expression of cell-cycle regulators, strains quickly accumulate mutations to get rid of the deleterious construction. Another limitation on the bacterial side can be the balance between envelope stability and permeability, which could limit the enlargement of bacteroids. The phenotype of *yej* mutants could be reflecting this problem. Beside their apparent increased sensitivity to membrane stress, they seem to push TBD further though not being able to survive it. Increasing envelope stability by tweaking their lipid, sterol or peptide importer composition could overcome this problem. If the limits over TBD is not imposed by the bacteria, it could be due to plant-driven constraints. The volume of plant cells could be restraining bacteroid development as being physically limited by the cell wall. Moreover, diffusibility of nitrogen in nodule tissues could be limiting. Nitrogen concentration is estimated through modelling to be around 112 μM in nodules and it is unsure

whether this concentration could be limiting for nitrogen fixation (Hunt et al., 1988). However, as shown by results with some bacterial cell-cycle mutant (unpublished data), trying to increase/accelerate the bacterial endoreduplication seem to lead to a substantial gain of biomass for the plant (up 50% dry weight gain compared to WT). Those hypothetical constraints might therefore not be reached yet if really existing.

vii. Conclusive remarks

Nitrogen fixation as applied to agriculture is of major interest. Indeed, it constitutes a viable alternative to fertilizer usage. Research on this field could help a transition towards a more sustainable agriculture. The studies performed during my thesis could help understanding the mechanisms underlying this symbiotic interaction. Indeed, TBD seems to be a leverage of choice to increase symbiotic efficiency and hopefully agricultural yields. This work helped to unveil new functions on *M. truncatula* NCR resistance of *S. meliloti* 1021 as well as deepening knowledge on already available mutants, by linking them to TBD and NCR peptides. These resistance mechanisms are not particularly specific of rhizobia, but seem necessary in various other biotic interactions, pathogenic included. This work could thus help to understand how bacteria manage to resist host defense under the form of AMP.

USDA110 and ORS285 symbiotic interactions with *A. afraspera* balanced the actual beliefs that USDA110 is undifferentiated. Though displaying unmodified morphology, the bacteria displays TBD-like features suggesting base functions might respond to the presence of NCR, even though this strain has not evolved with NCR-producing plants. Moreover, this work points out the necessity to develop data integration and prediction tools as a way to process the masses of data that accumulates. In the era of big data this computational work will become more and more essential, in order to make the most of the produced data and to be able to go from observation to prediction.

Aggregation of these findings with the current and future knowledge will be of major importance on the understanding of the functioning of the symbiotic process. They will hopefully help transform the agricultural techniques to overcome the problems linked to fertilizers usage. Three different strategies were proposed before and listed in the introduction to bypass fertilizers. One of them is already in use today, which is the addition of plant endophytic bacteria that fix nitrogen in crop cultures such as in seed coating techniques. The remaining two are more challenging and require much more knowledge and

engineering. Transferring the ability to plants themselves by expressing nitrogenase in specialized engineered organelles is one of them. However, it might require a deep understanding on how an organelle is entangled with its host cell. The second strategy, which would eventually be enhanced with the first, requires to transfer the ability to form symbiotic root nodules to cultivated plants such as cereals. Though cereals can be mycorrhized, deep modifications of the plant might be required to provide the ability to develop nodules. Managing to develop these strategies will likely be of major importance to palliate hunger in the world and pollution arising directly and indirectly from the Haber-Bosh process.

REFERENCES

- Aalto, A., Viitasaari, L., Ilmonen, P., Mombaerts, L., & Gonçalves, J. (2020). Gene regulatory network inference from sparsely sampled noisy data. *Nature Communications*, 2020. <https://doi.org/10.1038/s41467-020-17217-1>
- Abel, S., Abel zur Wiesch, P., Davis, B. M., & Waldor, M. K. (2015). Analysis of Bottlenecks in Experimental Models of Infection. *PLoS Pathogens*, 11(6), 1–7. <https://doi.org/10.1371/journal.ppat.1004823>
- Alazard, D. (1985). Stem and Root Nodulation in *Aeschynomene* spp. *Applied and Environmental Microbiology*, 50(3), 732–734. <https://doi.org/10.1128/aem.50.3.732-734.1985>
- Alazard, D. (1990). Nitrogen fixation in pure culture by rhizobia isolated from stem nodules of tropical *Aeschynomene* species. *FEMS Microbiology Letters*, 68(1–2), 177–182. [https://doi.org/10.1016/0378-1097\(90\)90147-I](https://doi.org/10.1016/0378-1097(90)90147-I)
- Alunni, B., & Gourion, B. (2016). Terminal bacteroid differentiation in the legume-rhizobium symbiosis: nodule-specific cysteine-rich peptides and beyond. *The New Phytologist*, 211(2), 411–417. <https://doi.org/10.1111/nph.14025>
- Alunni, B., Kevei, Z., Redondo-Nieto, M., Kondorosi, A., Mergaert, P., & Kondorosi, E. (2007). Genomic organization and evolutionary insights on GRP and NCR genes, two large nodule-specific gene families in *Medicago truncatula*. *Molecular Plant-Microbe Interactions*, 20(9), 1138–1148. <https://doi.org/10.1094/MPMI-20-9-1138>
- Alunni, B., & Mergaert, P. (2017). Défis et perspectives pour l'utilisation de la fixation biologique de l'azote en agriculture. In *Les sols et la vie souterraine, des enjeux majeurs en agroécologie*. (pp. 276–297).
- Alvarez-Buylla, E. R., Benítez, M., Dávila, E. B., Chaos, Á., Espinosa-Soto, C., & Padilla-Longoria, P. (2007). Gene regulatory network models for plant development. *Current Opinion in Plant Biology*, 10(1), 83–91. <https://doi.org/10.1016/j.pbi.2006.11.008>
- Amann, R. I., Ludwig, W., & Schleifer, K.-H. (1995). Phylogenetic Identification and In Situ Detection of Individual Microbial Cells without Cultivation. *MICROBIOLOGICAL REVIEWS*, 59(1), 143–169.
- Ané, J.-M., Kiss, G. B., Riely, B. K., Penmetsa, R. V., Oldroyd, G. E. D., Ayax, C., Lévy, J., Debelle, F., Baek, J. M., Kalo, P., Rosenberg, C., Roe, B. A., Long, S. R., Dénarié, J., & Cook, D. R. (2004). *Medicago truncatula* DMI1 Required for Bacterial and Fungal Symbioses in Legumes. *Science*, 303(5662), 1364–1367. <https://doi.org/10.1126/science.1092986>
- Ardisson, S., Kobayashi, H., Kambara, K., Rummel, C., Noel, K. D., Walker, G. C., Broughton, W. J., & Deakin, W. J. (2011). Role of BacA in lipopolysaccharide synthesis, peptide transport, and nodulation by *Rhizobium* sp. strain NGR234. *Journal of Bacteriology*, 193(9), 2218–2228. <https://doi.org/10.1128/JB.01260-10>
- Arnold, M. F. F., Caro-Hernandez, P., Tan, K., Runti, G., Wehmeier, S., Scocchi, M., Doerrler, W. T., Walker, G. C., & Ferguson, G. P. (2014). Enteric YaiW is a surface-exposed outer membrane lipoprotein that affects sensitivity to an antimicrobial peptide. *Journal of Bacteriology*, 196(2), 436–444. <https://doi.org/10.1128/JB.01179-13>
- Arnold, M. F. F., Shabab, M., Penterman, J., Boehme, K. L., Griffiths, J. S., & Walker, C. (2017). Genome-Wide Sensitivity Analysis of the Microsymbiont *Sinorhizobium meliloti* to Symbiotically Important, Defensin-Like Host Peptides. *MBio*, 8(4).

- Ausmees, N., Jonsson, H., Höglund, S., Ljunggren, H., & Lindberg, M. (1999). Structural and putative regulatory genes involved in cellulose synthesis in *Rhizobium leguminosarum* bv. trifolii. *Microbiology*, *145*(5), 1253–1262. <https://doi.org/10.1099/13500872-145-5-1253>
- Azani, N., Babineau, M., Bailey, C. D., Banks, H., Barbosa, A. R., Pinto, R. B., Boatwright, J. S., Borges, L. M., Brown, G. K., Bruneau, A., Candido, E., Cardoso, D., Chung, K. F., Clark, R. P., Conceição, A. D. S., Crisp, M., Cubas, P., Delgado-Salinas, A., Dexter, K. G., ... Zimmerman, E. (2017). A new subfamily classification of the leguminosae based on a taxonomically comprehensive phylogeny. *Taxon*, *66*(1), 44–77. <https://doi.org/10.12705/661.3>
- Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(25), 6506–6511. <https://doi.org/10.1073/pnas.1711842115>
- Barnett, M. J., Bittner, A. N., Toman, C. J., Oke, V., & Long, S. R. (2012). Dual RpoH sigma factors and transcriptional plasticity in a symbiotic bacterium. *Journal of Bacteriology*, *194*(18), 4983–4994. <https://doi.org/10.1128/JB.00449-12>
- Barrière, Q., Guefrachi, I., Gully, D., Lamouche, F., Pierre, O., Fardoux, J., Chaintreuil, C., Alunni, B., Timchenko, T., Giraud, E., & Mergaert, P. (2017). Integrated roles of BclA and DD-carboxypeptidase 1 in Bradyrhizobium differentiation within NCR-producing and NCR-lacking root nodules. *Scientific Reports*, *7*(1), 1–13. <https://doi.org/10.1038/s41598-017-08830-0>
- Baym, M., Lieberman, T. D., Kelsic, E. D., Chait, R., Gross, R., Yelin, I., & Kishony, R. (2016). Spatiotemporal microbial evolution on antibiotic landscapes. *Science*, *353*(6304), 1147–1152.
- Becquart-de Kozak, I., Reuhs, B. L., Buffard, D., Breda, C., Kim, J. S., Esnault, R., & Kondorosi, A. (1997). Role of the K-antigen subgroup of capsular polysaccharides in the early recognition process between *Rhizobium meliloti* and alfalfa leaves. *Molecular Plant-Microbe Interactions*, *10*(1), 114–123. <https://doi.org/10.1094/MPMI.1997.10.1.114>
- Beijerinck, M. W. (1888). Die bacterien der papilionaceenknöllchen. *Botanische Zeitung*, *46*(46), 725–735. https://scholar.google.com/scholar_lookup?title=W.%3ADie+Bakterien+der+Papilionaceenknöllchen&author=M.+Beijerinck&journal=Bot.+Ztg.+46%2C+725+%281888%29&volume=48&pages=837&publication_year=1890#d=gs_cit&u=%2Fscholar%3Fq%3Dinfo%3Atu2QFfbvDNSJ%3Ascholar
- Biondi, E. G., Reisinger, S. J., Skerker, J. M., Arif, M., Perchuk, B. S., Ryan, K. R., & Laub, M. T. (2006). Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature*, *444*(7121), 899–904. <https://doi.org/10.1038/nature05321>
- Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., Goormachtig, S., & Giraud, E. (2011). Nodulation of *Aeschynomene afraspera* and *A. indica* by Photosynthetic *Bradyrhizobium* Sp. Strain ORS285: The Nod-Dependent Versus the Nod-Independent Symbiotic Interaction. *Molecular Plant-Microbe Interactions*, *24*(11), 1359–1371. <https://doi.org/10.1094/MPMI-04-11-0093>
- Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., Goormachtig, S., Giraud, E., Tropicales, S., Ird, U. M. R., Inra, S., & Cirad, U. M. (2011). *Nodulation of Aeschynomene afraspera and A. indica by Photosynthetic Bradyrhizobium Sp. Strain ORS285: The Nod-Dependent Versus the Nod-Independent Symbiotic Interaction.* *24*(11), 1359–1371. <https://doi.org/10.1094/MPMI-04-11-0093>
- Bourassa, D. V., Kannenberg, E. L., Sherrier, D. J., Buhr, R. J., & Carlson, R. W. (2017). The lipopolysaccharide lipid a long-chain fatty acid is important for *Rhizobium leguminosarum* growth and stress adaptation in free-living and nodule environments. *Molecular Plant-Microbe Interactions*, *30*(2), 161–175. <https://doi.org/10.1094/MPMI-11-16-0230-R>

- Braga, L. P. P., Soucy, S. M., Amgarten, D. E., da Silva, A. M., & Setubal, J. C. (2018). Bacterial diversification in the light of the interactions with phages: The genetic symbionts and their role in ecological speciation. *Frontiers in Ecology and Evolution*, *6*(JAN), 0–12. <https://doi.org/10.3389/fevo.2018.00006>
- Breedveld, M. W., Canter Cremers, H. C. J., Batley, M., Posthumus, M. A., Zevenhuizen, L. P. T. M., Wijffelman, C. A., & Zehnder, A. J. B. (1993). Polysaccharide synthesis in relation to nodulation behavior of *Rhizobium leguminosarum*. *Journal of Bacteriology*, *175*(3), 750–757. <https://doi.org/10.1128/jb.175.3.750-757.1993>
- Breedveld, Michaël W., & Miller, K. J. (1994). Cyclic β -glucans of members of the family Rhizobiaceae. *Microbiological Reviews*, *58*(2), 145–161. <https://doi.org/10.1128/mmbr.58.2.145-161.1994>
- Brooks, J. F., Gyllborg, M. C., Cronin, D. C., Quillin, S. J., Mallama, C. A., Foxall, R., Whistler, C., Goodman, A. L., & Mandel, M. J. (2014). Global discovery of colonization determinants in the squid symbiont *Vibrio fischeri*. *Proceedings of the National Academy of Sciences*, *111*(48), 17284–17289. <https://doi.org/10.1073/pnas.1415957111>
- Broughton, W. J., Jabbouri, S., & Perret, X. (2000). Keys to symbiotic harmony. *Journal of Bacteriology*, *182*(20), 5641–5652. <https://doi.org/10.1128/JB.182.20.5641-5652.2000>
- Brown, T. A., Jones, M. K., Powell, W., & Allaby, R. G. (2009). The complex origins of domesticated crops in the Fertile Crescent. *Trends in Ecology and Evolution*, *24*(2), 103–109. <https://doi.org/10.1016/j.tree.2008.09.008>
- Bruneau, A., Doyle, J. J., Herendeen, P., Hughes, C., Kenicer, G., Lewis, G., Mackinder, B., Pennington, R. T., Sanderson, M. J., & Wojciechowski, M. F. (2013). Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon*, *62*(2), 217–248. <https://doi.org/10.12705/622.8>
- Bulgarelli, D., Schlaeppli, K., Spaepen, S., Van Themaat, E. V. L., & Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annual Review of Plant Biology*, *64*, 807–838. <https://doi.org/10.1146/annurev-arplant-050312-120106>
- Burill, T. J., & Hansen, R. (1917). Is symbiosis possible between legume bacteria and non-legume plants? *University of Illinois*, *202*.
- Busset, N., Lorenzo, F. Di, Palmigiano, A., Sturiale, L., Gressent, F., Fardoux, J., Gully, D., Chaintreuil, C., Molinaro, A., Silipo, A., & Giraud, E. (2017). The very long chain fatty acid (C26:25OH) linked to the lipid a is important for the fitness of the photosynthetic Bradyrhizobium strain ORS278 and the establishment of a successful symbiosis with *Aeschynomene* legumes. *Frontiers in Microbiology*, *8*(SEP), 1–13. <https://doi.org/10.3389/fmicb.2017.01821>
- Byrd, A. L., Belkaid, Y., & Segre, J. A. (2018). The human skin microbiome. *Nature Reviews Microbiology*, *16*(3), 143–155. <https://doi.org/10.1038/nrmicro.2017.157>
- Campbell, G. R. O., Reuhs, B. L., & Walker, G. C. (2002). Chronic intracellular infection of alfalfa nodules by *Sinorhizobium meliloti* requires correct lipopolysaccharide core. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(6), 3938–3943. <https://doi.org/10.1073/pnas.062425699>
- Cao, P., Lu, C., & Yu, Z. (2018). Historical nitrogen fertilizer use in agricultural ecosystems of the contiguous United States during 1850–2015: Application rate, timing, and fertilizer types. *Earth System Science Data*, *10*(2), 969–984. <https://doi.org/10.5194/essd-10-969-2018>
- Capela, D., Filipe, C., Bobik, C., Batut, J., & Bruand, C. (2006). *Sinorhizobium meliloti* differentiation during

- symbiosis with alfalfa: A transcriptomic dissection. *Molecular Plant-Microbe Interactions*, 19(4), 363–372. <https://doi.org/10.1094/MPMI-19-0363>
- Capoen, W., Sun, J., Wysham, D., Otegui, M. S., Venkateshwaran, M., Hirsch, S., Miwa, H., Downie, J. A., Morris, R. J., Ané, J.-M., & Oldroyd, G. E. D. (2011). Nuclear membranes control symbiotic calcium signaling of legumes. *Proceedings of the National Academy of Sciences of the United States of America*, 108(34), 14348–14353. <https://doi.org/10.1073/pnas.1107912108>
- Carro, L., Pujic, P., Alloisio, N., Fournier, P., Boubakri, H., Hay, A. E., Poly, F., François, P., Hocher, V., Mergaert, P., Balmand, S., Rey, M., Heddi, A., & Normand, P. (2015). Alnus peptides modify membrane porosity and induce the release of nitrogen-rich metabolites from nitrogen-fixing Frankia. *ISME Journal*, 9(8), 1723–1733. <https://doi.org/10.1038/ismej.2014.257>
- Cebolla, A., Vinardell, J. M., Kiss, E., Oláh, B., Roudier, F., Kondorosi, A., & Kondorosi, E. (1999). The mitotic inhibitor ccs52 is required for endoreduplication and ploidy-dependent cell enlargement in plants. *EMBO Journal*, 18(16), 4476–4484. <https://doi.org/10.1093/emboj/18.16.4476>
- Cescutti, P. (2010). Bacterial Capsular Polysaccharides and Exopolysaccharides. In *Microbial Glycobiology* (First edit). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-374546-0.00006-7>
- Chaintreuil, C., Arrighi, J. F., Giraud, E., Miché, L., Moulin, L., Dreyfus, B., Munive-Hernández, J. A., Villegas-Hernandez, M. D. C., & Béna, G. (2013). Evolution of symbiosis in the legume genus Aeschynomene. *New Phytologist*, 200(4), 1247–1259. <https://doi.org/10.1111/nph.12424>
- Chao, M. C., Abel, S., Davis, B. M., & Waldor, M. K. (2016). The design and analysis of transposon insertion sequencing experiments. *Nature Reviews Microbiology*, 14(2), 119–128. <https://doi.org/10.1038/nrmicro.2015.7>
- Charpentier, M., Sun, J., Martins, T. V., Radhakrishnan, G. V., Findlay, K., Soumpourou, E., Thouin, J., Véry, A., Sanders, D., Morris, R. J., & Oldroyd, G. E. D. (2016). Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science*, 352(6289), 1102–1105. <https://doi.org/10.1126/science.aae0109>
- Cheng, H. P., & Walker, G. C. (1998). Succinoglycan Is Required for Initiation and Elongation of Infection Threads during Nodulation of Alfalfa by Rhizobium meliloti. *J Bacteriol*, 180(19), 5183–5191. file:///C:/PDF/ChengWalker1998.pdf%0Ahttp://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10894742&dopt=Abstract
- Cheung, L. K. Y., Dupuis, J. H., Dee, D. R., Bryksa, B. C., & Yada, R. Y. (2020). Roles of Plant-Specific Inserts in Plant Defense. *Trends in Plant Science*, 25(7), 682–694. <https://doi.org/10.1016/j.tplants.2020.02.009>
- Chiu, C. H., & Paszkowski, U. (2020). Receptor-like kinases sustain symbiotic scrutiny. *Plant Physiology*, 182(4), 1597–1612. <https://doi.org/10.1104/PP.19.01341>
- Ciccarelli, F. D., Doerks, T., Von Mering, C., Creevey, C. J., Snel, B., & Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. *Science*, 311(5765), 1283–1287. <https://doi.org/10.1126/science.1123061>
- Ciocchini, A. E., Guidolin, L. S., Casabuono, A. C., Couto, A. S., Iñón De Iannino, N., & Ugalde, R. A. (2007). A glycosyltransferase with a length-controlling activity as a mechanism to regulate the size of polysaccharides. *Proceedings of the National Academy of Sciences of the United States of America*, 104(42), 16492–16497. <https://doi.org/10.1073/pnas.0708025104>
- Compant, S., Samad, A., Faist, H., & Sessitsch, A. (2019). A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *Journal of Advanced Research*, 19, 29–37.

<https://doi.org/10.1016/j.jare.2019.03.004>

- Corbalan, N., Runti, G., Adler, C., Covaceuszach, S., Ford, R. C., Lamba, D., Beis, K., Scocchi, M., & Vincent, P. A. (2013). Functional and structural study of the dimeric inner membrane protein SbmA. *Journal of Bacteriology*, *195*(23), 5352–5361. <https://doi.org/10.1128/JB.00824-13>
- Cosme, A. M., Becker, A., Santos, M. R., Sharypova, L. A., Santos, P. M., & Moreira, L. M. (2008). The outer membrane protein TolC from *Sinorhizobium meliloti* affects protein secretion, polysaccharide biosynthesis, antimicrobial resistance, and symbiosis. *Molecular Plant-Microbe Interactions*, *21*(7), 947–957. <https://doi.org/10.1094/MPMI-21-7-0947>
- Covert, M. W. (2009). Integrated Regulatory and Metabolic Models. In *Computational Systems Biology*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-088786-6.50029-0>
- Crespo-Rivas, J. C., Guefrachi, I., Mok, K. C., Villaécija-Aguilar, J. A., Acosta-Jurado, S., Pierre, O., Ruiz-Sainz, J. E., Taga, M. E., Mergaert, P., & Vinardell, J. M. (2016). *Sinorhizobium fredii* HH103 bacteroids are not terminally differentiated and show altered O-antigen in nodules of the Inverted Repeat-Lacking Clade legume *Glycyrrhiza uralensis*. *Environmental Microbiology*, *18*(8), 2392–2404. <https://doi.org/10.1111/1462-2920.13101>
- Crespo-Rivas, J. C., Margaret, I., Hidalgo, A., Buendia-Claveria, A. M., Ollero, F. J., Lopez-Baena, F. J., Del Socorro Murdoch, P., Rodriguez-Carvajal, M. A., Soria-Diaz, 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., & 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. *Molecular Plant-Microbe Interactions*, *22*(5), 575–588. <https://doi.org/10.1094/MPMI-22-5-0575>
- Cress, B. F., Englaender, J. A., He, W., Kasper, D., Linhardt, R. J., & Koffas, M. A. G. (2014). Masquerading microbial pathogens: Capsular polysaccharides mimic host-tissue molecules. *FEMS Microbiology Reviews*, *38*(4), 660–697. <https://doi.org/10.1111/1574-6976.12056>
- 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., & Giraud, E. (2015). Convergent evolution of endosymbiont differentiation in dalbergioid and inverted repeat-lacking clade legumes mediated by nodule-specific cysteine-rich peptides. *Plant Physiology*, *169*(2), 1254–1265. <https://doi.org/10.1104/pp.15.00584>
- D’Haeze, W., & Holsters, M. (2002). Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology*, *12*(6), 79–105. <https://doi.org/10.1093/glycob/12.6.79R>
- Daniel J. Gage. (2004). Infection and Invasion of Roots by Symbiotic, Nitrogen-Fixing Rhizobia during Nodulation of Temperate Legumes. *Microbiology and Molecular Biology Reviews: MMBR*, *68*(June), 203. <https://doi.org/10.1128/MMBR.68.2.280>
- Davidson, A. L., Dassa, E., Orelle, C., & Chen, J. (2008). Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiology and Molecular Biology Reviews*, *72*(2), 317–364. <https://doi.org/10.1128/mubr.00031-07>
- Delgado, M. J., Bedmar, E. J., & Downie, J. A. (1998). Genes involved in the formation and assembly of rhizobial cytochromes and their role in symbiotic nitrogen fixation. *Advances in Microbial Physiology*, *40*, 191–231. [https://doi.org/10.1016/s0065-2911\(08\)60132-0](https://doi.org/10.1016/s0065-2911(08)60132-0)
- Delory, M., Hallez, R., Letesson, J. J., & De Bolle, X. (2006). An RpoH-like heat shock sigma factor is involved in stress response and virulence in *Brucella melitensis* 16M. *Journal of Bacteriology*, *188*(21), 7707–7710. <https://doi.org/10.1128/JB.00644-06>

- Delwiche, C. F. (1999). Tracing the thread of plastid diversity through the tapestry of life. *American Naturalist*, *154*(4 SUPPL.). <https://doi.org/10.2307/2463984>
- Dharmatilake, A. J., & Bauer, W. D. (1992). Chemotaxis of *Rhizobium meliloti* towards nodulation gene-inducing compounds from alfalfa roots. *Applied and Environmental Microbiology*, *58*(4), 1153–1158. <https://doi.org/10.1128/aem.58.4.1153-1158.1992>
- DiCenzo, G. C., Checcucci, A., Bazzicalupo, M., Mengoni, A., Viti, C., Dziejewit, L., Finan, T. M., Galardini, M., & Fondi, M. (2016). Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. *Nature Communications*, *7*. <https://doi.org/10.1038/ncomms12219>
- DiCenzo, G. C., Tesi, M., Pfau, T., Mengoni, A., & Fondi, M. (2019). A Virtual Nodule Environment (ViNE) for modelling the inter-kingdom metabolic 2 integration during symbiotic nitrogen fixation. *BioArchives*.
- diCenzo, G., Milunovic, B., Cheng, J., & Finan, T. M. (2013). The tRNA^{arg} gene and engA Are essential genes on the 1.7-Mb pSymb megaplasmid of *sinorhizobium meliloti* and were translocated together from the chromosome in an ancestral strain. *Journal of Bacteriology*, *195*(2), 202–212. <https://doi.org/10.1128/JB.01758-12>
- Dobrindt, U., Hochhut, B., Hentschel, U., & Hacker, J. (2004). Genomic islands in pathogenic and environmental microorganisms. *Nature Reviews Microbiology*, *2*(5), 414–424. <https://doi.org/10.1038/nrmicro884>
- Domenech, P., Kobayashi, H., Levier, K., Walker, G. C., & Barry, C. E. (2009). BacA, an ABC transporter involved in maintenance of chronic murine infections with mycobacterium tuberculosis. *Journal of Bacteriology*, *191*(2), 477–485. <https://doi.org/10.1128/JB.01132-08>
- Dreyfus, B., Garcia, J. L., & Gillis, M. (1988). Characterization of *Azorhizobium caulinodans* gen. nov. sp. nov., a Stem-Nodulating Nitrogen-Fixing Bacterium Isolated from *Sesbania rostrata*. *International Journal of Systematic Bacteriology*, *98*(1), 89–98.
- Dupont, L., Alloing, G., Pierre, O., El, S., Hopkins, J., Hrouart, D., & Frenedo, P. (2012). The Legume Root Nodule: From Symbiotic Nitrogen Fixation to Senescence. *Senescence*. <https://doi.org/10.5772/34438>
- Durgo, H., Klement, E., Hunyadi-Gulyas, E., Szucs, A., Kereszt, A., Medzihradzsky, K. F., & Kondorosi, E. (2015). Identification of nodule-specific cysteine-rich plant peptides in endosymbiotic bacteria. *Proteomics*, *15*(13), 2291–2295. <https://doi.org/10.1002/pmic.201400385>
- Erickson, H. P., Anderson, D. E., & Osawa, M. (2010). FtsZ in Bacterial Cytokinesis: Cytoskeleton and Force Generator All in One. *Microbiology and Molecular Biology Reviews*, *74*(4), 504–528. <https://doi.org/10.1128/mubr.00021-10>
- Eswarappa, S. M., Panguluri, K. K., Hensel, M., & Chakravorty, D. (2008). The yejABEF operon of *Salmonella* confers resistance to antimicrobial peptides and contributes to its virulence. *Microbiology*, *154*(2), 666–678. <https://doi.org/10.1099/mic.0.2007/011114-0>
- Fabian, B. K., Tetu, S. G., & Paulsen, I. T. (2020). Application of Transposon Insertion Sequencing to Agricultural Science. *Frontiers in Plant Science*, *11*(March). <https://doi.org/10.3389/fpls.2020.00291>
- Fabre, S., Gully, D., Poitout, A., Patrel, D., Arrighi, J. F., Giraud, E., Czernic, P., & Cartieaux, F. (2015). Nod factor-independent nodulation in *Aeschynomene evenia* required the common plant-microbe symbiotic toolkit. *Plant Physiology*, *169*(4), 2654–2664. <https://doi.org/10.1104/pp.15.01134>
- Fan, Y., & Pedersen, O. (2020). Gut microbiota in human metabolic health and disease. *Nature Reviews*

Microbiology. <https://doi.org/10.1038/s41579-020-0433-9>

- Fang, G., Passalacqua, K. D., Hocking, J., Llopis, P. M., Gerstein, M., Bergman, N. H., & Jacobs-Wagner, C. (2013). Transcriptomic and phylogenetic analysis of a bacterial cell cycle reveals strong associations between gene co-expression and evolution. *BMC Genomics*, *14*(1), 1. <https://doi.org/10.1186/1471-2164-14-450>
- Faria, J. P., Overbeek, R., Taylor, R. C., Conrad, N., Vonstein, V., Goelzer, A., Fromion, V., Rocha, M., Rocha, I., & Henry, C. S. (2016). Reconstruction of the regulatory network for *Bacillus subtilis* and reconciliation with gene expression data. *Frontiers in Microbiology*, *7*(MAR), 1–11. <https://doi.org/10.3389/fmicb.2016.00275>
- Farkas, A., Maróti, G., Dürgo, H., Györgypál, Z., Lima, R. M., Medzihradsky, K. F., Kereszt, A., Mergaert, P., & Kondorosi, É. (2014). Medicago truncatula symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(14), 5183–5188. <https://doi.org/10.1073/pnas.1404169111>
- Ferguson, G. P., Dattat, A., Baumgartner, J., Roop, R. M., Carlson, R. W., & Walker, G. C. (2004). Similarity to peroxisomal-membrane protein family reveals that *Sinorhizobium* and *Brucella* BacA affect lipid-A fatty acids. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(14), 5012–5017. <https://doi.org/10.1073/pnas.0307137101>
- Ferguson, G. P., Roop, R. M., & Walker, G. C. (2002). Deficiency of a *Sinorhizobium meliloti* bacA mutant in alfalfa symbiosis correlates with alteration of the cell envelope. *Journal of Bacteriology*, *184*(20), 5625–5632. <https://doi.org/10.1128/JB.184.20.5625-5632.2002>
- Fernández-López, M., Goormachtig, S., Gao, M., D’Haeze, W., Van Montagu, M., & Holsters, M. (1998). Ethylene-mediated phenotypic plasticity in root nodule development on *Sesbania rostrata*. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(21), 12724–12728. <https://doi.org/10.1073/pnas.95.21.12724>
- Fernández, D., Toffanin, A., Palacios, J. M., Ruiz-Argüeso, T., & Imperial, J. (2005). Hydrogenase genes are uncommon and highly conserved in *Rhizobium leguminosarum* bv. *viciae*. *FEMS Microbiology Letters*, *253*(1), 83–88. <https://doi.org/10.1016/j.femsle.2005.09.022>
- Fields, S. (2004). Global Nitrogen: Cycling out of Control. *Environmental Health Perspectives*, *112*(10), 556–563.
- Fonouni-Farde, C., Tan, S., Baudin, M., Brault, M., Wen, J., Mysore, K. S., Niebel, A., Frugier, F., & Diet, A. (2016). DELLA-mediated gibberellin signalling regulates Nod factor signalling and rhizobial infection. *Nature Communications*, *7*. <https://doi.org/10.1038/ncomms12636>
- Fournier, J., Timmers, A. C. J., Sieberer, B. J., Jauneau, A., Chabaud, M., & Barker, D. G. (2008). Mechanism of infection thread elongation in root hairs of *Medicago truncatula* and dynamic interplay with associated rhizobial colonization. *Plant Physiology*, *148*(4), 1985–1995. <https://doi.org/10.1104/pp.108.125674>
- Fraysse, N., Lindner, B., Kaczynski, Z., Sharypova, L., Holst, O., Niehaus, K., & Poinso, V. (2005). *Sinorhizobium meliloti* strain 1021 produces a low-molecular-mass capsular polysaccharide that is a homopolymer of 3-deoxy-D-manno-oct-2-ulosonic acid harboring a phospholipid anchor. *Glycobiology*, *15*(1), 101–108. <https://doi.org/10.1093/glycob/cwh142>
- Fraysse, Nicolas, Couderc, F., & Poinso, V. (2003). Surface polysaccharide involvement in establishing the rhizobium-legume symbiosis. *European Journal of Biochemistry*, *270*(7), 1365–1380. <https://doi.org/10.1046/j.1432-1033.2003.03492.x>

- Futahashi, R., Tanaka, K., Tanahashi, M., Nikoh, N., Kikuchi, Y., Lee, B. L., & Fukatsu, T. (2013). Gene Expression in Gut Symbiotic Organ of Stinkbug Affected by Extracellular Bacterial Symbiont. *PLoS ONE*, *8*(5). <https://doi.org/10.1371/journal.pone.0064557>
- Gage, D. J. (2002). Analysis of Infection Thread Development Using Gfp- and DsRed-Expressing *Sinorhizobium meliloti*. *JOURNAL OF BACTERIOLOGY*, *184*(24), 7042–7046. <https://doi.org/10.1128/JB.184.24.7042>
- Galibert, F., Finan, T. M., Long, S. R., Pühler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M. J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R. W., ... Batut, J. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science*, *293*(5530), 668–672. <https://doi.org/10.1126/science.1060966>
- Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., Bakkoury, M. El, Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., ... Johnston, M. (2002). Functional profiling using the *Saccharomyces* genome. *Nature*, *418*, 387–391. <https://doi.org/10.1101/pdb.prot088039>
- Gibson, K. E., Kobayashi, H., & Walker, G. C. (2008). Molecular determinants of a symbiotic chronic infection. *Annual Review of Genetics*, *42*(170), 413–441. <https://doi.org/10.1146/annurev.genet.42.110807.091427>
- 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., ... Sadowsky, M. (2007). Legumes Symbioses: Absence of Nod Genes in Photosynthetic Bradyrhizobia. *Science*, *316*(5829). <http://science.sciencemag.org/content/316/5829/1307>
- Glazebrook, J., Ichige, A., & Walker, G. C. (1993). A *Rhizobium meliloti* homolog of the *Escherichia coli* peptide-antibiotic transport protein SbmA is essential for bacteroid development. *Genes and Development*, *7*(8), 1485–1497. <https://doi.org/10.1101/gad.7.8.1485>
- González-Sama, A., De La Peña, T. C., Kevei, Z., Mergaert, P., Lucas, M. M., De Felipe, M. R., Kondorosi, E., & Pueyo, J. J. (2006). Nuclear DNA endoreduplication and expression of the mitotic inhibitor Ccs52 associated to determinate and lupinoid nodule organogenesis. *Molecular Plant-Microbe Interactions*, *19*(2), 173–180. <https://doi.org/10.1094/MPMI-19-0173>
- Goormachtig, S., Mergaert, P., Van Montagu, M., & Holsters, M. (1998). The symbiotic interaction between *Azorhizobium caulinodans* and *Sesbania rostrata* molecular cross-talk in a beneficial plant-bacterium interaction. *Sub-Cellular Biochemistry*, *29*, 117–164. https://doi.org/10.1007/978-1-4899-1707-2_4
- Goormachtig, Sofie, Capoen, W., & Holsters, M. (2004). *Rhizobium* infection: Lessons from the versatile nodulation behaviour of water-tolerant legumes. *Trends in Plant Science*, *9*(11), 518–522. <https://doi.org/10.1016/j.tplants.2004.09.005>
- Gourion, B., & Alunni, B. (2018). Strain-specific symbiotic genes: A new level of control in the intracellular accommodation of *Rhizobia* within legume nodule cells. *Molecular Plant-Microbe Interactions*, *31*(3), 287–288. <https://doi.org/10.1094/MPMI-01-18-0010-LE>
- Gourion, B., Berrabah, F., Ratet, P., & Stacey, G. (2015). *Rhizobium*-legume symbioses: The crucial role of plant immunity. *Trends in Plant Science*, *20*(3), 186–194. <https://doi.org/10.1016/j.tplants.2014.11.008>
- Gruber, N., & Galloway, J. N. (2008). An Earth-system perspective of the global nitrogen cycle. *Nature*, *451*(7176), 293–296. <https://doi.org/10.1038/nature06592>

- Guefrachi, I., Pierre, O., Timchenko, T., Barrière, Q., Czernic, P., Verly, C., Mars, M., Kondorosi, E., Giraud, E., & Mergaert, P. (2015). *Bradyrhizobium BclA Is a Peptide Transporter Required for Bacterial Differentiation in Symbiosis with Aeschynomene Legumes*. *28*(11), 1155–1166.
- Gully, D., Gargani, D., Bonaldi, K., Grangeteau, C., Chaintreuil, C., Fardoux, J., Nguyen, P., Marchetti, R., Nouwen, N., Molinaro, A., Mergaert, P., & Giraud, E. (2016). A Peptidoglycan-remodeling enzyme is critical for bacteroid differentiation in *Bradyrhizobium* spp. during legume symbiosis. *Molecular Plant-Microbe Interactions*, *29*(6), 447–457. <https://doi.org/10.1094/MPMI-03-16-0052-R>
- Haag, A. F., Baloban, M., Sani, M., Kerscher, B., Pierre, O., Farkas, A., Longhi, R., Boncompagni, E., Hérouart, D., Dall'Angelo, S., Kondorosi, E., Zanda, M., Mergaert, P., & Ferguson, G. P. (2011). Protection of sinorhizobium against host cysteine-rich antimicrobial peptides is critical for symbiosis. *PLoS Biology*, *9*(10). <https://doi.org/10.1371/journal.pbio.1001169>
- Haag, A. F., Kerscher, B., Dall'Angelo, S., Sani, M., Longhi, R., Baloban, M., Wilson, H. M., Mergaert, P., Zanda, M., & Ferguson, G. P. (2012). Role of cysteine residues and disulfide bonds in the activity of a legume root nodule-specific, cysteine-rich peptide. *Journal of Biological Chemistry*, *287*(14), 10791–10798. <https://doi.org/10.1074/jbc.M111.311316>
- Haag, A. F., Wehmeier, S., Beck, S., Marlow, V. L., Fletcher, V., James, E. K., & Ferguson, G. P. (2009). The *Sinorhizobium meliloti* LpxXL and AcpXL proteins play important roles in bacteroid development within alfalfa. *Journal of Bacteriology*, *191*(14), 4681–4686. <https://doi.org/10.1128/JB.00318-09>
- Hacker, S., Gödeke, J., Lindemann, A., Mesa, S., Pessi, G., & Narberhaus, F. (2008). Global consequences of phosphatidylcholine reduction in *Bradyrhizobium japonicum*. *Molecular Genetics and Genomics*, *280*(1), 59–72. <https://doi.org/10.1007/s00438-008-0345-2>
- Hakoyama, T., Niimi, K., Watanabe, H., Tabata, R., Matsubara, J., Sato, S., Nakamura, Y., Tabata, S., Jichun, L., Matsumoto, T., Tatsumi, K., Nomura, M., Tajima, S., Ishizaka, M., Yano, K., Imaizumi-Anraku, H., Kawaguchi, M., Kouchi, H., & Sugauma, N. (2009). Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. *Nature*, *462*(7272), 514–517. <https://doi.org/10.1038/nature08594>
- Hall, N. (2007). Advanced sequencing technologies and their wider impact in microbiology. *Journal of Experimental Biology*, *210*(9), 1518–1525. <https://doi.org/10.1242/jeb.001370>
- Hansen, A. K., & Moran, N. A. (2011). Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 2849–2854. <https://doi.org/10.1073/pnas.1013465108>
- Haskett, T. L., Terpolilli, J. J., Bekuma, A., O'Hara, G. W., Sullivan, J. T., Wang, P., Ronson, C. W., & Ramsay, J. P. (2016). Assembly and transfer of tripartite integrative and conjugative genetic elements. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(43), 12268–12273. <https://doi.org/10.1073/pnas.1613358113>
- Hauser, F., Pessi, G., Friberg, M., Weber, C., Rusca, N., Lindemann, A., Fischer, H. M., & Hennecke, H. (2007). Dissection of the *Bradyrhizobium japonicum* NifA+ σ 54 regulon, and identification of a ferredoxin gene (fdxN) for symbiotic nitrogen fixation. *Molecular Genetics and Genomics*, *278*(3), 255–271. <https://doi.org/10.1007/s00438-007-0246-9>
- Haygood, M. G., & Distel, D. L. (1993). Bioluminescent symbionts of flashlight fishes and deep-sea anglerfishes form unique lineages related to the genus *Vibrio*. *Nature*, *363*(6425), 154–156. <https://doi.org/10.1038/363154a0>
- Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S. N., Richelle, A., Heinken, A., Haraldsdóttir, H. S., Wachowiak, J., Keating, S. M., Vlasov, V., Magnúsdóttir, S., Ng, C. Y., Preciat, G., Žagare, A., Chan, S. H. J., Aurich,

- M. K., Clancy, C. M., Modamio, J., Sauls, J. T., ... Fleming, R. M. T. (2019). Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nature Protocols*, *14*(3), 639–702. <https://doi.org/10.1038/s41596-018-0098-2>
- Hews, C. L., Cho, T., Rowley, G., & Raivio, T. L. (2019). Maintaining Integrity Under Stress: Envelope Stress Response Regulation of Pathogenesis in Gram-Negative Bacteria. *Frontiers in Cellular and Infection Microbiology*, *9*(September), 1–25. <https://doi.org/10.3389/fcimb.2019.00313>
- Hirsch, A. M. (1992). Developmental Biology of Legume Nodulation. *New Phytologist*, *122*(40), 211–237.
- Horrigan, L., Lawrence, R. S., & Walker, P. (2002). How sustainable agriculture can address the environmental and human health harms of industrial agriculture. *Environmental Health Perspectives*, *110*(5), 445–456. <https://doi.org/10.1289/ehp.02110445>
- Horvath, B., Domonkos, A., Kereszt, A., Szucs, A., Abrahám, E., Ayaydin, F., Boka, K., Chen, Y., Chen, R., Murray, J. D., Udvardi, M. K., Kondorosi, E., & Kalo, P. (2015). Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the *Medicago truncatula* *dnf7* mutant. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(49), 15232–15237. <https://doi.org/10.1073/pnas.1500777112>
- Huddleston, J. R. (2014). Horizontal gene transfer in the human gastrointestinal tract: Potential spread of antibiotic resistance genes. *Infection and Drug Resistance*, *7*, 167–176. <https://doi.org/10.2147/IDR.S48820>
- Hunt, S., Gaito, S. T., & Layzell, D. B. (1988). Model of gas exchange and diffusion in legume nodules. *Planta*, *173*(1), 128–141. <https://doi.org/10.1007/bf00394497>
- Ichige, A., & Walker, G. C. (1997). Genetic analysis of the *Rhizobium meliloti* *bacA* gene: Functional interchangeability with the *Escherichia coli* *sbmA* gene and phenotypes of mutants. *Journal of Bacteriology*, *179*(1), 209–216. <https://doi.org/10.1128/jb.179.1.209-216.1997>
- Iida, T., Itakura, M., Anda, M., Sugawara, M., Isawa, T., Okubo, T., Sato, S., Chiba-Kakizaki, K., & Minamisawa, K. (2015). Symbiosis island shuffling with abundant insertion sequences in the genomes of extra-slow-growing strains of soybean bradyrhizobia. *Applied and Environmental Microbiology*, *81*(12), 4143–4154. <https://doi.org/10.1128/AEM.00741-15>
- Itoh, H., Aita, M., Nagayama, A., Meng, X., Kamagata, Y., Navarro, R., Hori, T., Ohgiya, S., & Kikuchi, Y. (2014). Evidence of environmental and vertical transmission of Burkholderia symbionts in the oriental chinch bug, *Cavelerius saccharivorus* (Heteroptera: Blissidae). *Applied and Environmental Microbiology*, *80*(19), 5974–5983. <https://doi.org/10.1128/AEM.01087-14>
- Jiao, Y. S., Liu, Y. H., Yan, H., Wang, E. T., Tian, C. F., Chen, W. X., Guo, B. L., & Chen, W. F. (2015). Rhizobial diversity and nodulation characteristics of the extremely promiscuous legume *sophora flavescens*. *Molecular Plant-Microbe Interactions*, *28*(12), 1338–1352. <https://doi.org/10.1094/MPMI-06-15-0141-R>
- Johnston, C., Martin, B., Fichant, G., Polard, P., & Claverys, J. P. (2014). Bacterial transformation: Distribution, shared mechanisms and divergent control. *Nature Reviews Microbiology*, *12*(3), 181–196. <https://doi.org/10.1038/nrmicro3199>
- Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., & Walker, G. C. (2007). How rhizobial symbionts invade plants: The *Sinorhizobium* - *Medicago* model. *Nature Reviews Microbiology*, *5*(8), 619–633. <https://doi.org/10.1038/nrmicro1705>
- Kaiser, B. N., Moreau, S., Castelli, J., Thomson, R., Lambert, A., Bogliolo, S., Puppo, A., & Day, D. A. (2003). The soybean NRAMP homologue, *GmDMT1*, is a symbiotic divalent metal transporter capable of

- ferrous iron transport. *Plant Journal*, 35(3), 295–304. <https://doi.org/10.1046/j.1365-313X.2003.01802.x>
- Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., ... Tabata, S. (2000). Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Research*, 7(6), 331–338. <https://doi.org/10.1093/dnares/7.6.331>
- Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpo, S., Tsuruoka, H., Wada, T., Yamada, M., & Tabata, S. (2002). Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Research*, 9(6), 189–197. <https://doi.org/10.1093/dnares/9.6.189>
- Kannenbergh, E. L., Perzl, M., Mtiller, P., Härtner, T., & Poralla, K. (1996). *Hopanoid lipids in Bradyrhizobium and other plant-associated bacteria and cloning of the Bradyrhizobium japonicum squalene-hopene cyclase gene. i*, 107–112.
- Karunakaran, R., Haag, A. F., East, A. K., Ramachandran, V. K., Prell, J., James, E. K., Scocchi, M., Ferguson, G. P., & Poole, P. S. (2010). BacA is essential for bacteroid development in nodules of galeoid, but not phaseoloid, legumes. *Journal of Bacteriology*, 192(11), 2920–2928. <https://doi.org/10.1128/JB.00020-10>
- Karygianni, L., Ren, Z., Koo, H., & Thurnheer, T. (2020). Biofilm Matrixome: Extracellular Components in Structured Microbial Communities. *Trends in Microbiology*, 28(8), 668–681. <https://doi.org/10.1016/j.tim.2020.03.016>
- Kawaharada, Y., Kelly, S., Nielsen, M. W., Hjuler, C. T., Gysel, K., Muszyński, A., Carlson, R. W., Thygesen, M. B., Sandal, N., Asmussen, M. H., Vinther, M., Andersen, S. U., Krusell, L., Thirup, S., Jensen, K. J., Ronson, C. W., Blaise, M., Radutoiu, S., & Stougaard, J. (2015). Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature*, 523(7560), 308–312. <https://doi.org/10.1038/nature14611>
- Kazmierczak, T., Nagymihaly, M., Lamouche, F., Barrière, Q., Guefrachi, I., Alunni, B., Ouadghiri, M., Ibijbjen, J., Kondorosi, É., Mergaert, P., & Gruber, V. (2017). Specific host-responsive associations between medicago truncatula accessions and sinorhizobium strains. *Molecular Plant-Microbe Interactions*, 30(5), 399–409. <https://doi.org/10.1094/MPMI-01-17-0009-R>
- Kereszt, A., Kiss, E., Reuhs, B. L., Carlson, R. W., Kondorosi, Á., & Putnoky, P. (1998). Novel rkp gene clusters of *Sinorhizobium meliloti* involved in capsular polysaccharide production and invasion of the symbiotic nodule: The rkpK gene encodes a UDP-glucose dehydrogenase. *Journal of Bacteriology*, 180(20), 5426–5431. <https://doi.org/10.1128/jb.180.20.5426-5431.1998>
- Kereszt, A., Mergaert, P., Montiel, J., Endre, G., & Kondorosi, É. (2018). Impact of plant peptides on symbiotic nodule development and functioning. *Frontiers in Plant Science*, 9(July), 1–16. <https://doi.org/10.3389/fpls.2018.01026>
- Kevei, Z., Loughnon, G., Mergaert, P., Horváth, G. V., Kereszt, A., Jayaraman, D., Zaman, N., Marcel, F., Regulski, K., Kiss, G. B., Kondorosi, A., Endre, G., Kondorosi, E., & Ané, J.-M. (2007). 3-Hydroxy-3-methylglutaryl coenzyme A reductase1 interacts with NOR1 and is crucial for nodulation in *Medicago truncatula*. *Plant Cell*, 19(12), 3974–3989. <https://doi.org/10.1105/tpc.107.053975>
- Kevei, Z., Vinardell, J. M., Kiss, G. B., Kondorosi, A., & Kondorosi, E. (2002). Glycine-rich proteins encoded by a nodule-specific gene family are implicated in different stages of symbiotic nodule development in *Medicago* Spp. *Molecular Plant-Microbe Interactions*, 15(9), 922–931.

<https://doi.org/10.1094/MPMI.2002.15.9.922>

- Kiers, E. T., Hutton, M. G., & Denison, R. F. (2007). Human selection and the relaxation of legume defences against ineffective rhizobia. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1629), 3119–3126. <https://doi.org/10.1098/rspb.2007.1187>
- Kikuchi, Y., Hayatsu, M., Hosokawa, T., Nagayama, A., Tago, K., & Fukatsu, T. (2012). Symbiont-mediated insecticide resistance. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(22), 8618–8622. <https://doi.org/10.1073/pnas.1200231109>
- Kikuchi, Y., Hosokawa, T., & Fukatsu, T. (2007). Insect-microbe mutualism without vertical transmission: A stinkbug acquires a beneficial gut symbiont from the environment every generation. *Applied and Environmental Microbiology*, *73*(13), 4308–4316. <https://doi.org/10.1128/AEM.00067-07>
- Kim, J. K., Won, Y. J., Nikoh, N., Nakayama, H., Han, S. H., Kikuchi, Y., Rhee, Y. H., Park, H. Y., Kwon, J. Y., Kurokawa, K., Dohmae, N., Fukatsu, T., & Lee, B. L. (2013). Polyester synthesis genes associated with stress resistance are involved in an insect-bacterium symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(26). <https://doi.org/10.1073/pnas.1303228110>
- Kim, M., Chen, Y., Xi, J., Waters, C., Chen, R., & Wang, D. (2015). An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(49), 15238–15243. <https://doi.org/10.1073/pnas.1500123112>
- Kim, S., Zeng, W., Bernard, S., Liao, J., Venkateshwaran, M., Ané, J.-M., & Jiang, Y. (2019). Ca²⁺-regulated Ca²⁺ channels with an RCK gating ring control plant symbiotic associations. *Nature Communications*, *10*(1), 1–12. <https://doi.org/10.1038/s41467-019-11698-5>
- Koga, R., Meng, X. Y., Tsuchida, T., & Fukatsu, T. (2012). Cellular mechanism for selective vertical transmission of an obligate insect symbiont at the bacteriocyte-embryo interface. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(20), 1230–1237. <https://doi.org/10.1073/pnas.1119212109>
- Kondorosi, E., Mergaert, P., & Kereszt, A. (2013). A Paradigm for endosymbiotic life: Cell differentiation of rhizobium bacteria provoked by host plant factors. *Annual Review of Microbiology*, *67*, 611–628. <https://doi.org/10.1146/annurev-micro-092412-155630>
- Krusell, L., Krause, K., Ott, T., Desbrosses, G., Krämer, U., Sato, S., Nakamura, Y., Tabata, S., James, E. K., Sandal, N., Stougaard, J., Kawaguchi, M., Miyamoto, A., Suganuma, N., & Udvardi, M. K. (2005). The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. *Plant Cell*, *17*(5), 1625–1636. <https://doi.org/10.1105/tpc.104.030106>
- Krysciak, D., Grote, J., Orbegoso, M. R., Utpatel, C., Förstner, K. U., Li, L., Schmeisser, C., Krishnan, H. B., & Streit, W. R. (2014). RNA sequencing analysis of the broad-host-range strain *Sinorhizobium fredii* NGR234 identifies a large set of genes linked to quorum sensing-dependent regulation in the background of a *traI* and *ngri* deletion mutant. *Applied and Environmental Microbiology*, *80*(18), 5655–5671. <https://doi.org/10.1128/AEM.01835-14>
- Kulkarni, G., Busset, N., Molinaro, A., Gargani, D., Chaintreuil, C., Silipo, A., Giraud, E., & Newman, D. K. (2015). Specific hopanoid classes differentially affect free-living and symbiotic states of bradyrhizobium diazoefficiens. *MBio*, *6*(5), 1–9. <https://doi.org/10.1128/mBio.01251-15>
- Lamouche, F., Bonadé-Bottino, N., Mergaert, P., & Alunni, B. (2019). Symbiotic efficiency of spherical and elongated bacteroids in the *Aeschynomene*-bradyrhizobium symbiosis. *Frontiers in Plant Science*, *10*(April). <https://doi.org/10.3389/fpls.2019.00377>

- Lamouche, F., Gully, D., Chaumeret, A., Nouwen, N., Verly, C., Pierre, O., Sciallano, C., Fardoux, J., Jeudy, C., Szücs, A., Mondy, S., Salon, C., Nagy, I., Kereszt, A., Dessaux, Y., Giraud, E., Mergaert, P., & Alunni, B. (2019). Transcriptomic dissection of Bradyrhizobium sp. strain ORS285 in symbiosis with Aeschynomene spp. inducing different bacteroid morphotypes with contrasted symbiotic efficiency. *Environmental Microbiology*, *21*(9), 3244–3258. <https://doi.org/10.1111/1462-2920.14292>
- Lancaster, K. M., Roemelt, M., Ettenhuber, P., Hu, Y., Ribbe, M. W., Neese, F., Bergmann, U., & DeBeer, S. (2011). X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. *Science*, *334*(6058), 974–977. <https://doi.org/10.1126/science.1206445>
- Lang, J., & Faure, D. (2014). Functions and regulation of quorum-sensing in *Agrobacterium tumefaciens*. *Frontiers in Plant Science*, *5*(JAN), 1–13. <https://doi.org/10.3389/fpls.2014.00014>
- Le Quéré, A. J. L., Deakin, W. J., Schmeisser, C., Carlson, R. W., Streit, W. R., Broughton, W. J., & Forsberg, L. S. (2006). Structural characterization of a K-antigen capsular polysaccharide essential for normal symbiotic infection in *Rhizobium* sp. NGR234: Deletion of the rkpMNO locus prevents synthesis of 5,7-diacetamido-3,5,7,9-tetra-deoxy-non-2-ulosonic acid. *Journal of Biological Chemistry*, *281*(39), 28981–28992. <https://doi.org/10.1074/jbc.M513639200>
- Lehman, A. P., & Long, S. R. (2013). Exopolysaccharides from *Sinorhizobium meliloti* can protect against H₂O₂-dependent damage. *Journal of Bacteriology*, *195*(23), 5362–5369. <https://doi.org/10.1128/JB.00681-13>
- Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J. C., & Dénarié, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature*, *344*(6268), 781–784. <https://doi.org/10.1038/344781a0>
- Letellier, L., & Bonhivers, M. (1996). Chapter 27 intrinsic and extrinsic channels in bacteria. *Handbook of Biological Physics*, *2*(C), 615–636. [https://doi.org/10.1016/S1383-8121\(96\)80068-6](https://doi.org/10.1016/S1383-8121(96)80068-6)
- Leung, C. W. T., Hong, Y., Zhao, J. H., Chen, S., Pletneva, E. V., & Tang, B. Z. (2013). Superior fluorescent probe for detection of potassium ion. *Talanta*, *86*, 1263–1268. <https://doi.org/10.1016/j.talanta.2015.06.015>
- LeVier, K., Phillips, R. W., Grippe, V. K., Roop, R. M., & Walker, G. C. (2000). Similar requirements of a plant symbiont and a mammalian pathogen for prolonged intracellular survival. *Science*, *287*(5462), 2492–2493. <https://doi.org/10.1126/science.287.5462.2492>
- LeVier, Kristin, & Walker, G. C. (2001). Genetic analysis of the *Sinorhizobium meliloti* BacA protein: Differential effects of mutations on phenotypes. *Journal of Bacteriology*, *183*(21), 6444–6453. <https://doi.org/10.1128/JB.183.21.6444-6453.2001>
- 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., & Debelle, F. (2004). A Putative Ca²⁺ and Calmodulin-Dependent Protein Kinase Required for Bacterial and Fungal Symbioses. *Science (New York, N.Y.)*, *303*(5662), 1361–1364. <http://www.ncbi.nlm.nih.gov/pubmed/14963335>
- Lewis, G., Schrire, B., MacKinder, B., & Lock, M. (2005). Legumes of the World. *Edinburgh Journal of Botany*, *62*(3), 195–196. <https://doi.org/10.1017/s0960428606190198>
- Limpens, E., Mirabella, R., Fedorova, E., Franken, C., Franssen, H., Bisseling, T., & Geurts, R. (2005). Formation of organelle-like N₂-fixing symbiosomes in legume root nodules is controlled by DMI2. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(29), 10375–10380. <https://doi.org/10.1073/pnas.0504284102>

- Loh, J., & Stacey, G. (2003). Nodulation Gene Regulation in *Bradyrhizobium japonicum*: a Unique Integration of Global Regulatory Circuits. *Society*, *69*(1), 10–17. <https://doi.org/10.1128/AEM.69.1.10>
- Lopez, D., & Koch, G. (2017). Exploring functional membrane microdomains in bacteria: an overview. *Current Opinion in Microbiology*, *36*, 76–84. <https://doi.org/10.1016/j.mib.2017.02.001>
- Lupp, C., & Ruby, E. G. (2005). *Vibrio fischeri* uses two quorum-sensing systems for the regulation of early and late colonization factors. *Journal of Bacteriology*, *187*(11), 3620–3629. <https://doi.org/10.1128/JB.187.11.3620-3629.2005>
- Madigan, M. T., & Martinko, J. M. (2006). *Brock Biology of Microorganisms*. Pearson Prentice Hall. <https://books.google.fr/books?id=O7DuAAAAMAAJ>
- Madsen, L. H., Tirichine, L., Jurkiewicz, A., Sullivan, J. T., Heckmann, A. B., Bek, A. S., Ronson, C. W., James, E. K., & Stougaard, J. (2010). The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nature Communications*, *1*(1). <https://doi.org/10.1038/ncomms1009>
- Maillet, F., Poinso, V., André, O., Puech-Pagés, V., Haouy, A., Gueunier, M., Cromer, L., Giraudet, D., Formey, D., Niebel, A., Martinez, E. A., Driguez, H., Bécard, G., & Dénarié, J. (2011). Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. *Nature*, *469*(7328), 58–64. <https://doi.org/10.1038/nature09622>
- Maldonado, R. F., Sá-Correia, I., & Valvano, M. A. (2016). Lipopolysaccharide modification in gram-negative bacteria during chronic infection. *FEMS Microbiology Reviews*, *40*(4), 480–493. <https://doi.org/10.1093/femsre/fuw007>
- Marchal, K., & Vanderleyden, J. (2000). The “oxygen paradox” of dinitrogen-fixing bacteria. *Biology and Fertility of Soils*, *30*(5–6), 363–373. <https://doi.org/10.1007/s003740050017>
- Mardirossian, M., Grzela, R., Giglione, C., Meinel, T., Gennaro, R., Mergaert, P., & Scocchi, M. (2014). The host antimicrobial peptide Bac71-35 binds to bacterial ribosomal proteins and inhibits protein synthesis. *Chemistry and Biology*, *21*(12), 1639–1647. <https://doi.org/10.1016/j.chembiol.2014.10.009>
- Margaret, I., Becker, A., Blom, J., Bonilla, I., Goesmann, A., Göttfert, M., Lloret, J., Mittard-Runte, V., Rückert, C., Ruiz-Sainz, J. E., Vinardell, J. M., & Weidner, S. (2011). Symbiotic properties and first analyses of the genomic sequence of the fast growing model strain *Sinorhizobium fredii* HH103 nodulating soybean. *Journal of Biotechnology*, *155*(1), 11–19. <https://doi.org/10.1016/j.jbiotec.2011.03.016>
- Marlow, V. L., Haag, A. F., Kobayashi, H., Fletcher, V., Scocchi, M., Walker, G. C., & Ferguson, G. P. (2009). Essential role for the BacA protein in the uptake of a truncated eukaryotic peptide in *Sinorhizobium meliloti*. *Journal of Bacteriology*, *191*(5), 1519–1527. <https://doi.org/10.1128/JB.01661-08>
- Maruya, J., & Saeki, K. (2010). The bacA Gene Homolog, mlr7400, in *Mesorhizobium loti* MAFF303099 is dispensable for symbiosis with *Lotus japonicus* but partially capable of supporting the symbiotic function of bacA in *Sinorhizobium meliloti*. *Plant and Cell Physiology*, *51*(9), 1443–1452. <https://doi.org/10.1093/pcp/pcq114>
- Massaiu, I., Pasotti, L., Sonnenschein, N., Rama, E., Cavaletti, M., Magni, P., Calvio, C., & Herrgård, M. J. (2019). Integration of enzymatic data in *Bacillus subtilis* genome-scale metabolic model improves phenotype predictions and enables in silico design of poly- γ -glutamic acid production strains. *Microbial Cell Factories*, *18*(1), 1–20. <https://doi.org/10.1186/s12934-018-1052-2>
- Masson-Boivin, C., Giraud, E., Perret, X., & Batut, J. (2009). Establishing nitrogen-fixing symbiosis with

- legumes: how many rhizobium recipes? *Trends in Microbiology*, 17(10), 458–466. <https://doi.org/10.1016/j.tim.2009.07.004>
- Masuda, S., Hennecke, H., & Fischer, H. M. (2017). Requirements for efficient thiosulfate oxidation in *Bradyrhizobium diazoefficiens*. *Genes*, 8(12), 1–12. <https://doi.org/10.3390/genes8120390>
- Masuda, S., Saito, M., Sugawara, C., Itakura, M., Eda, S., & Minamisawa, K. (2016). Identification of the hydrogen uptake gene cluster for chemolithoautotrophic growth and symbiosis hydrogen uptake in *Bradyrhizobium Diazoefficiens*. *Microbes and Environments*, 31(1), 76–78. <https://doi.org/10.1264/j sme2.ME15182>
- Matthysse, A. G., & McMahan, S. (1998). Root colonization by *Agrobacterium tumefaciens* is reduced in *cel*, *attB*, *attD*, and *attR* mutants. *Applied and Environmental Microbiology*, 64(7), 2341–2345. <https://doi.org/10.1128/aem.64.7.2341-2345.1998>
- Mattiuzzo, M., Bandiera, A., Gennaro, R., Benincasa, M., Pacor, S., Antcheva, N., & Scocchi, M. (2007). Role of the *Escherichia coli* *SbmA* in the antimicrobial activity of proline-rich peptides. *Molecular Microbiology*, 66(1), 151–163. <https://doi.org/10.1111/j.1365-2958.2007.05903.x>
- McFadden, G. I. (2014). Origin and evolution of plastids and photosynthesis in eukaryotes. *Cold Spring Harbor Perspectives in Biology*, 6(4), 1–9. <https://doi.org/10.1101/cshperspect.a016105>
- McFall-Ngai, M. (2014). Divining the Essence of Symbiosis: Insights from the Squid-Vibrio Model. *PLoS Biology*, 12(2), 1–6. <https://doi.org/10.1371/journal.pbio.1001783>
- Mergaert, P. (2018). Role of antimicrobial peptides in controlling symbiotic bacterial populations. *Natural Product Reports*, 35(4), 336–356. <https://doi.org/10.1039/c7np00056a>
- Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., & 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.*
- Mergaert, Peter, Kikuchi, Y., Shigenobu, S., & Nowack, E. C. M. (2017). Metabolic Integration of Bacterial Endosymbionts through Antimicrobial Peptides. *Trends in Microbiology*, 25(9), 703–712. <https://doi.org/10.1016/j.tim.2017.04.007>
- Mergaert, Peter, Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.-E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., & Kondorosi, E. (2006). *Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium – legume symbiosis*.
- Mergaert, Peter, Van Montagu, M., & Holsters, M. (1997). Molecular mechanisms of Nod factor diversity. *Molecular Microbiology*, 25(5), 811–817. <https://doi.org/10.1111/j.1365-2958.1997.mmi526.x>
- Mikuláss, K. R., Nagy, K., Bogos, B., Szegletes, Z., Kovács, E., Farkas, A., Váró, G., Kondorosi, É., & Kereszt, A. (2016). Antimicrobial nodule-specific cysteine-rich peptides disturb the integrity of bacterial outer and inner membranes and cause loss of membrane potential. *Annals of Clinical Microbiology and Antimicrobials*, 15(1), 1–5. <https://doi.org/10.1186/s12941-016-0159-8>
- Mileykovskaya, E., & Dowhan, W. (2009). Cardiolipin membrane domains in prokaryotes and eukaryotes. *Biochimica et Biophysica Acta - Biomembranes*, 1788(10), 2084–2091. <https://doi.org/10.1016/j.bbamem.2009.04.003>
- Miller, A. J., & Cramer, M. D. (2004). Root nitrogen acquisition and assimilation. In *Plant and Soil* (Vol. 274, Issues 1–2). <https://doi.org/10.1007/s11104-004-0965-1>
- Minchin, F. R., James, E. K., & Becana, M. (2008). Oxygen Diffusion, Production Of Reactive Oxygen And Nitrogen Species, And Antioxidants In Legume Nodules. *Nitrogen-Fixing Leguminous Symbioses*,

- 321–362. https://doi.org/10.1007/978-1-4020-3548-7_11
- Minder, A. C., De Rudder, K. E. E., Narberhaus, F., Fischer, H. M., Hennecke, H., & Geiger, O. (2001). Phosphatidylcholine levels in Bradyrhizobium japonicum membranes are critical for an efficient symbiosis with the soybean host plant. *Molecular Microbiology*, *39*(5), 1186–1198. <https://doi.org/10.1046/j.1365-2958.2001.02325.x>
- Mitsui, H., Sato, T., Sato, Y., Ito, N., & Minamisawa, K. (2004). Sinorhizobium meliloti RpoH1 is required for effective nitrogen-fixing symbiosis with alfalfa. *Molecular Genetics and Genomics*, *271*(4), 416–425. <https://doi.org/10.1007/s00438-004-0992-x>
- Montiel, J., Downie, J. A., Farkas, A., Bihari, P., Herczeg, R., Bálint, B., Mergaert, P., Kereszt, A., & Kondorosi, É. (2017). Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(19), 5041–5046. <https://doi.org/10.1073/pnas.1704217114>
- Montiel, J., Szucs, A., Boboescu, I. Z., Gherman, V. D., Kondorosi, É., & Kereszt, A. (2016). Terminal bacteroid differentiation is associated with variable morphological changes in legume species belonging to the inverted repeat-lacking clade. *Molecular Plant-Microbe Interactions*, *29*(3), 210–219. <https://doi.org/10.1094/MPMI-09-15-0213-R>
- Mora, Y., Díaz, R., Vargas-Lagunas, C., Peralta, H., Guerrero, G., Aguilar, A., Encarnación, S., Girard, L., & Mora, J. (2014). Nitrogen-fixing rhizobial strains isolated from common bean seeds: Phylogeny, physiology, and genome analysis. *Applied and Environmental Microbiology*, *80*(18), 5644–5654. <https://doi.org/10.1128/AEM.01491-14>
- Moreau, S., Thomson, R. M., Kaiser, B. N., Trevaskis, B., Guerinot, M. Lou, Udvardi, M. K., Puppo, A., & Day, D. A. (2002). GmZIP1 encodes a symbiosis-specific zinc transporter in soybean. *Journal of Biological Chemistry*, *277*(7), 4738–4746. <https://doi.org/10.1074/jbc.M106754200>
- Morris, B. E. L., Henneberger, R., Huber, H., & Moissl-Eichinger, C. (2013). Microbial syntrophy: Interaction for the common good. *FEMS Microbiology Reviews*, *37*(3), 384–406. <https://doi.org/10.1111/1574-6976.12019>
- Moulin, L., Munive, a, Dreyfus, B., & Boivin-Masson, C. (2001). Nodulation of legumes by members of the beta-subclass of Proteobacteria. *Nature*, *411*(June), 948–950. <https://doi.org/10.1038/35082070>
- Müller, A., Wenzel, M., Strahl, H., Grein, F., Saaki, T. N. V., & Kohl, B. (2016). *Daptomycin inhibits cell envelope synthesis by interfering with fluid membrane microdomains*. <https://doi.org/10.1073/pnas.1611173113>
- Musovic, S., Oregaard, G., Kroer, N., & Sørensen, S. J. (2006). Cultivation-independent examination of horizontal transfer and host range of an IncP-1 plasmid among gram-positive and gram-negative bacteria indigenous to the barley rhizosphere. *Applied and Environmental Microbiology*, *72*(10), 6687–6692. <https://doi.org/10.1128/AEM.00013-06>
- Nett, R. S., Montanares, M., Marcassa, A., Lu, X., Nagel, R., Charles, T. C., Hedden, P., Rojas, M. C., & Peters, R. J. (2017). Elucidation of gibberellin biosynthesis in bacteria reveals convergent evolution. *Nature Chemical Biology*, *13*(1), 69–74. <https://doi.org/10.1038/nchembio.2232>
- Newton, W. (1998). Nitrogénases: fonction et évolution. *Bulletin Société Française de Microbiologie*, 238–241.
- Nikaido, H. (2003). Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiology and Molecular Biology Reviews*, *67*(4), 593–656. <https://doi.org/10.1128/mmbr.67.4.593-656.2003>

- Norman, A., Hansen, L. H., & Sørensen, S. J. (2009). Conjugative plasmids: Vessels of the communal gene pool. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1527), 2275–2289. <https://doi.org/10.1098/rstb.2009.0037>
- Nouwen, N., Arrighi, J. F., Cartieaux, F., Chaintreuil, C., Gully, D., Klopp, C., & Giraud, E. (2017). The role of rhizobial (NifV) and plant (FEN1) homocitrate synthases in Aeschynomene/photosynthetic Bradyrhizobium symbiosis. *Scientific Reports*, *7*(1), 1–10. <https://doi.org/10.1038/s41598-017-00559-0>
- Novikova, M., Metlitskaya, A., Datsenko, K., Kazakov, T., Kazakov, A., Wanner, B., & Severinov, K. (2007). The Escherichia coli Yej transporter is required for the uptake of translation inhibitor microcin C. *Journal of Bacteriology*, *189*(22), 8361–8365. <https://doi.org/10.1128/JB.01028-07>
- Okazaki, S., Kaneko, T., Sato, S., & Saeki, K. (2013). Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 17131–17136. <https://doi.org/10.1073/pnas.1302360110>
- Okazaki, S., Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., Gully, D., Arrighi, J. F., & Furuta, N. (2016). *Rhizobium – legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS*. 64–74. <https://doi.org/10.1038/ismej.2015.103>
- Oldroyd, G. E. D. (2013). Speak, friend, and enter: Signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology*, *11*(4), 252–263. <https://doi.org/10.1038/nrmicro2990>
- Oldroyd, G. E. D., & Dixon, R. (2014). Biotechnological solutions to the nitrogen problem. *Current Opinion in Biotechnology*, *26*, 19–24. <https://doi.org/10.1016/j.copbio.2013.08.006>
- Oldroyd, G. E. D., Murray, J. D., Poole, P. S., & Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annual Review of Genetics*, *45*, 119–144. <https://doi.org/10.1146/annurev-genet-110410-132549>
- Oono, R., & Denison, R. F. (2010). Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiology*, *154*(3), 1541–1548. <https://doi.org/10.1104/pp.110.163436>
- Pecrix, Y., Staton, S. E., Sallet, E., Lelandais-Brière, C., Moreau, S., Carrère, S., Blein, T., Jardinaud, M. F., Latrasse, D., Zouine, M., Zahm, M., Kreplak, J., Mayjonade, B., Satgé, C., Perez, M., Cauet, S., Marande, W., Chantry-Darmon, C., Lopez-Roques, C., ... Gamas, P. (2018). Whole-genome landscape of Medicago truncatula symbiotic genes. *Nature Plants*, *4*(12), 1017–1025. <https://doi.org/10.1038/s41477-018-0286-7>
- Peng, J. (2019). Gene redundancy and gene compensation: An updated view. *Journal of Genetics and Genomics*, *46*(7), 329–333. <https://doi.org/10.1016/j.jgg.2019.07.001>
- Penterman, J., Abo, R. P., De Nisco, N. J., Arnold, M. F. F., Longhi, R., Zanda, M., & Walker, G. C. (2014). Host plant peptides elicit a transcriptional response to control the Sinorhizobium meliloti cell cycle during symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(9), 3561–3566. <https://doi.org/10.1073/pnas.1400450111>
- Perez-Galdona, R., & Kahn, M. L. (1994). Effects of organic acids and low pH on Rhizobium meliloti 104A14. *Microbiology*, *140*(5), 1231–1235. <https://doi.org/10.1099/13500872-140-5-1231>
- Peters, J. W., Stowell, M. H. B., Soltis, S. M., Finnegan, M. G., Johnson, M. K., & Rees, D. C. (1997). Redox-dependent structural changes in the nitrogenase P-cluster. *Biochemistry*, *36*(6), 1181–1187. <https://doi.org/10.1021/bi9626665>
- Peters, N. K., Frost, J. W., & Long, S. R. (1986). A plant flavone, luteolin, induces expression of Rhizobium

- meliloti nodulation genes. *Science*, *233*(4767), 977–980. <https://doi.org/10.1126/science.3738520>
- Phillips, D. A., & Tsai, S. M. (1992). Flavonoids as plant signals to rhizosphere microbes. *Mycorrhiza*, *1*(2), 55–58. <https://doi.org/10.1007/BF00206136>
- Pierre, O., Hopkins, J., Combiér, M., Baldacci, F., Engler, G., Brouquisse, R., Hérouart, D., & Boncompagni, E. (2014). Involvement of papain and legumain proteinase in the senescence process of *Medicago truncatula* nodules. *New Phytologist*, *202*(3), 849–863. <https://doi.org/10.1111/nph.12717>
- Pini, F., De Nisco, N. J., Ferri, L., Penterman, J., Fioravanti, A., Brillì, M., Mengoni, A., Bazzicalupo, M., Viollier, P. H., Walker, G. C., & Biondi, E. G. (2015). Cell Cycle Control by the Master Regulator CtrA in *Sinorhizobium meliloti*. *PLoS Genetics*, *11*(5), 1–24. <https://doi.org/10.1371/journal.pgen.1005232>
- Pletz, M. W. R., McGee, L., Van Beneden, C. A., Petit, S., Bardsley, M., Barlow, M., & Klugman, K. P. (2006). Fluoroquinolone resistance in invasive *Streptococcus pyogenes* isolates due to spontaneous mutation and horizontal gene transfer. *Antimicrobial Agents and Chemotherapy*, *50*(3), 943–948. <https://doi.org/10.1128/AAC.50.3.943-948.2006>
- Poole, P., Ramachandran, V., & Terpolilli, J. (2018). Rhizobia: From saprophytes to endosymbionts. *Nature Reviews Microbiology*, *16*(5), 291–303. <https://doi.org/10.1038/nrmicro.2017.171>
- Prell, J., White, J. P., Bourdes, A., Bunnewell, S., Bongaerts, R. J., & Poole, P. S. (2009). Legumes regulate *Rhizobium* bacteroid development and persistence by the supply of branched-chain amino acids. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(30), 12477–12482. <https://doi.org/10.1073/pnas.0903653106>
- Price, P. A., Tanner, H. R., Dillon, B. A., Shabab, M., Walker, G. C., & Griffitts, J. S. (2015). Rhizobial peptidase HrrP cleaves host-encoded signaling peptides and mediates symbiotic compatibility. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(49), 15244–15249. <https://doi.org/10.1073/pnas.1417797112>
- Pueppke, S. G., & Broughton, W. J. (1999). *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. *Molecular Plant-Microbe Interactions*, *12*(4), 293–318. <https://doi.org/10.1094/MPMI.1999.12.4.293>
- Ramos, B., Robledo, M., Jime, J. I., Dazzo, F., & Mateos, P. F. (2008). *Rhizobium cellulase CelC2 is essential for primary symbiotic infection of legume host roots*. *105*(19).
- Rastogi, R. P., Madamwar, D., & Incharoensakdi, A. (2015). Bloom dynamics of cyanobacteria and their toxins: Environmental health impacts and mitigation strategies. *Frontiers in Microbiology*, *6*(NOV), 1–22. <https://doi.org/10.3389/fmicb.2015.01254>
- Remigi, P., Zhu, J., Young, J. P. W., & Masson-Boivin, C. (2016). Symbiosis within Symbiosis: Evolving Nitrogen-Fixing Legume Symbionts. *Trends in Microbiology*, *24*(1), 63–75. <https://doi.org/10.1016/j.tim.2015.10.007>
- Renier, A., Maillet, F., Fardoux, J., Poinso, V., Giraud, E., & Nouwen, N. (2011). Photosynthetic *Bradyrhizobium* sp. strain ORS285 synthesizes 2-O-methylfucosylated lipochitooligosaccharides for nod gene-dependent interaction with *Aeschynomene* plants. *Molecular Plant-Microbe Interactions*, *24*(12), 1440–1447. <https://doi.org/10.1094/MPMI-05-11-0104>
- Ribeiro, C. W., Baldacci-Cresp, F., Pierre, O., Larousse, M., Benyamina, S., Lambert, A., Hopkins, J., Castella, C., Cazareth, J., Alloing, G., Boncompagni, E., Couturier, J., Mergaert, P., Gamas, P., Rouhier, N., Montrichard, F., & Frendo, P. (2017). Regulation of Differentiation of Nitrogen-Fixing Bacteria by Microsymbiont Targeting of Plant Thioredoxin s1. *Current Biology*, *27*(2), 250–256. <https://doi.org/10.1016/j.cub.2016.11.013>

- Roberts, D. M., & Tyerman, S. D. (2002). Voltage-dependent cation channels permeable to nh_4^+ , k^+ , and ca^{2+} in the symbiosome membrane of the model legume *lotus japonicus*. *Plant Physiology*, *128*(2), 370–378. <https://doi.org/10.1104/pp.010568>
- Rodríguez-Navarro, D. N., Rodríguez-Carvajal, M. A., Acosta-Jurado, S., Soto, M. J., Margaret, I., Crespo-Rivas, J. C., Sanjuan, J., Temprano, F., Gil-Serrano, A., Ruiz-Sainz, J. E., & Vinardell, J. M. (2014). Structure and biological roles of *Sinorhizobium fredii* HH103 exopolysaccharide. *PLoS ONE*, *9*(12), 1–31. <https://doi.org/10.1371/journal.pone.0115391>
- Rogel, M. A., Ormeño-Orrillo, E., & Martínez Romero, E. (2011). Symbiovars in rhizobia reflect bacterial adaptation to legumes. *Systematic and Applied Microbiology*, *34*(2), 96–104. <https://doi.org/10.1016/j.syapm.2010.11.015>
- Roger, A. J., Muñoz-Gómez, S. A., & Kamikawa, R. (2017). The Origin and Diversification of Mitochondria. *Current Biology*, *27*(21), R1177–R1192. <https://doi.org/10.1016/j.cub.2017.09.015>
- Rotundo, J. L., Borrás, L., de Bruin, J. D., & Pedersen, P. (2014). Soybean nitrogen uptake and utilization in Argentina and united states cultivars. *Crop Science*, *54*(3), 1153–1165. <https://doi.org/10.2135/cropsci2013.09.0618>
- Roux, B., Rodde, N., Jardinaud, M. F., Timmers, T., Sauviac, L., Cottret, L., Carrère, S., Sallet, E., Courcelle, E., Moreau, S., Debellé, F., Capela, D., De Carvalho-Niebel, F., Gouzy, J., Bruand, C., & Gamas, P. (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant Journal*, *77*(6), 817–837. <https://doi.org/10.1111/tpj.12442>
- Rowley, G., Spector, M., Kormanec, J., & Roberts, M. (2006). Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. *Nature Reviews. Microbiology*, *4*(5), 383–394. <https://doi.org/10.1038/nrmicro1394>
- Rubio, L. M., & Ludden, P. W. (2008). Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Annual Review of Microbiology*, *62*, 93–111. <https://doi.org/10.1146/annurev.micro.62.081307.162737>
- Runti, G., Lopez Ruiz, M. del C., Stoilova, T., Hussain, R., Jennions, M., Choudhury, H. G., Benincasa, M., Gennaro, R., Beis, K., & Scocchi, M. (2013). Functional characterization of SbmA, a bacterial inner membrane transporter required for importing the antimicrobial peptide Bac7(1-35). *Journal of Bacteriology*, *195*(23), 5343–5351. <https://doi.org/10.1128/JB.00818-13>
- Sachetto-Martins, G., Franco, L. O., & De Oliveira, D. E. (2000). Plant glycine-rich proteins: A family or just proteins with a common motif? *Biochimica et Biophysica Acta - Gene Structure and Expression*, *1492*(1), 1–14. [https://doi.org/10.1016/S0167-4781\(00\)00064-6](https://doi.org/10.1016/S0167-4781(00)00064-6)
- Sáenz, J. P., Sezgin, E., Schwille, P., & Simons, K. (2012). Functional convergence of hopanoids and sterols in membrane ordering. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(35), 14236–14240. <https://doi.org/10.1073/pnas.1212141109>
- Salton, M. R. J., & Kim, K.-S. (1996). Structure. *Medical Microbiology*, *4*. <https://doi.org/10.1515/9781942401698-006>
- Seefeldt, A. C., Graf, M., Pérébasquine, N., Nguyen, F., Arenz, S., Mardirossian, M., Scocchi, M., Wilson, D. N., & Innis, C. A. (2016). Structure of the mammalian antimicrobial peptide Bac7(1-16) bound within the exit tunnel of a bacterial ribosome. *Nucleic Acids Research*, *44*(5), 2429–2438. <https://doi.org/10.1093/nar/gkv1545>
- Seefeldt, L. C., Hoffman, B. M., & Dean, D. R. (2009). Mechanism of mo-dependent nitrogenase. *Annual*

- Review of Biochemistry*, 78, 701–722. <https://doi.org/10.1146/annurev.biochem.78.070907.103812>
- Sen, D., & Weaver, R. W. (1984). A basis for different rates of N₂-fixation by the same strains of *Rhizobium* in peanut and cowpea root nodules. *Plant Science Letters*, 34(3), 239–246. [https://doi.org/10.1016/S0304-4211\(84\)80002-4](https://doi.org/10.1016/S0304-4211(84)80002-4)
- Shabab, M., Arnold, M. F. F., Penterman, J., Wommack, A. J., Bocker, H. T., Price, P. A., Griffiths, J. S., Nolan, E. M., & Walker, G. C. (2016). Disulfide cross-linking influences symbiotic activities of nodule peptide NCR247. *Proceedings of the National Academy of Sciences of the United States of America*, 113(36), 10157–10162. <https://doi.org/10.1073/pnas.1610724113>
- Sharypova, L. A., Niehaus, K., Scheidle, H., Holst, O., & Becker, A. (2003). *Sinorhizobium meliloti* acpXL mutant lacks the C28 hydroxylated fatty acid moiety of lipid A and does not express a slow migrating form of lipopolysaccharide. *Journal of Biological Chemistry*, 278(15), 12946–12954. <https://doi.org/10.1074/jbc.M209389200>
- Shoemaker, N. B., Vlamakis, H., Hayes, K., & Salyers, A. A. (2001). Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. *Applied and Environmental Microbiology*, 67(2), 561–568. <https://doi.org/10.1128/AEM.67.2.561-568.2001>
- Silipo, A., Leone, M. R., Erbs, G., Lanzetta, R., Parrilli, M., Chang, W.-S., Newman, M.-A., & Molinaro, A. (2011). A Unique Bicyclic Monosaccharide from the Bradyrhizobium Lipopolysaccharide and Its Role in the Molecular Interaction with Plants. *Angewandte Chemie*, 123(52), 12818–12820. <https://doi.org/10.1002/ange.201106548>
- Silipo, A., Vitiello, G., Gully, D., Sturiale, L., Chaintreuil, C., Fardoux, J., Gargani, D., Lee, H.-I., Kulkarni, G., Busset, N., Marchetti, R., Palmigiano, A., Moll, H., Engel, R., Lanzetta, R., Paduano, L., Parrilli, M., Chang, W.-S., Holst, O., ... Molinaro, A. (2014). Covalently linked hopanoid-lipid A improves outer-membrane resistance of a Bradyrhizobium symbiont of legumes. *Nature Communications*, 5. <https://doi.org/10.1038/ncomms6106>
- Simonsen, A. K., DInnage, R., Barrett, L. G., Prober, S. M., & Thrall, P. H. (2017). Symbiosis limits establishment of legumes outside their native range at a global scale. *Nature Communications*, 8, 1–9. <https://doi.org/10.1038/ncomms14790>
- Simpson, F. B., & Burris, R. H. (1984). A nitrogen pressure of 50 atmospheres does not prevent evolution of hydrogen by nitrogenase. *Science*, 224(4653), 1095–1097. <https://doi.org/10.1126/science.6585956>
- Smil, V. (2011). Nitrogen cycle and world food production. *World Agriculture*, 2(Smil), 9–13. <http://www.vaclavsmil.com/wp-content/uploads/docs/smil-article-worldagriculture.pdf>
- Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C., & de la Cruz, F. (2010). Mobility of Plasmids. *Microbiology and Molecular Biology Reviews*, 74(3), 434–452. <https://doi.org/10.1128/mnbr.00020-10>
- Smit, G., Tubbing, D. M. J., Kijne, J. W., & Lugtenberg, B. J. J. (1991). Role of Ca²⁺ in the activity of rhicadhesin from *Rhizobium leguminosarum* biovar *viciae*, which mediates the first step in attachment of Rhizobiaceae cells to plant root hair tips. *Archives of Microbiology*, 155(3), 278–283. <https://doi.org/10.1007/BF00252212>
- Sniegowski, P., Lenski, R. E., & Gerrish, P. (1997). Evolution of high mutation rates in experimental populations of *E. coli*. *Nature*, 387(June), 703–705.
- Sprent, J. I., Ardley, J., & James, E. K. (2017). Biogeography of nodulated legumes and their nitrogen-

- fixing symbionts. *New Phytologist*, 215(1), 40–56. <https://doi.org/10.1111/nph.14474>
- Stanway, R. R., Bushell, E., Chiappino-Pepe, A., Roques, M., Sanderson, T., Franke-Fayard, B., Caldelari, R., Golomingi, M., Nyonda, M., Pandey, V., Schwach, F., Chevalley, S., Ramesar, J., Metcalf, T., Herd, C., Burda, P. C., Rayner, J. C., Soldati-Favre, D., Janse, C. J., ... Heussler, V. T. (2019). Genome-Scale Identification of Essential Metabolic Processes for Targeting the Plasmodium Liver Stage. *Cell*, 179(5), 1112–1128.e26. <https://doi.org/10.1016/j.cell.2019.10.030>
- Starker, C. G., Parra-Colmenares, A. L., Smith, L., Mitra, R. M., & Long, S. R. (2006). Nitrogen fixation mutants of *Medicago truncatula* fail to support plant and bacterial symbiotic gene expression. *Plant Physiology*, 140(2), 671–680. <https://doi.org/10.1104/pp.105.072132>
- Steimle, A., Autenrieth, I. B., & Frick, J. S. (2016). Structure and function: Lipid A modifications in commensals and pathogens. *International Journal of Medical Microbiology*, 306(5), 290–301. <https://doi.org/10.1016/j.ijmm.2016.03.001>
- Stephens, W. Z., Wiles, T. J., Martinez, E. S., Jemielita, M., Burns, A. R., Parthasarathy, R., Bohannan, B. J. M., & Guillemin, K. (2015). Identification of population bottlenecks and colonization factors during assembly of bacterial communities within the zebrafish intestine. *MBio*, 6(6), 1–11. <https://doi.org/10.1128/mBio.01163-15>
- Strullu-Derrien, C., Selosse, M. A., Kenrick, P., & Martin, F. M. (2018). The origin and evolution of mycorrhizal symbioses: from palaeomycology to phylogenomics. *New Phytologist*, 220(4), 1012–1030. <https://doi.org/10.1111/nph.15076>
- Sy, A., Giraud, E., Jourand, P., Garcia, N., Willems, A., De Lajudie, P., Prin, Y., Neyra, M., Gillis, M., Boivin-Masson, C., & Dreyfus, B. (2001). Methylo-trophic Methylobacterium bacteria nodulate and fix nitrogen in symbiosis with legumes. *Journal of Bacteriology*, 183(1), 214–220. <https://doi.org/10.1128/JB.183.1.214-220.2001>
- Szczyglowski, K., Kapranov, P., Hamburger, D., & De Bruijn, F. J. (1998). The *Lotus japonicus* LjNOD70 nodulin gene encodes a protein with similarities to transporters. *Plant Molecular Biology*, 37(4), 651–661. <https://doi.org/10.1023/A:1006043428636>
- Tamminen, M., Virta, M., Fani, R., & Fondi, M. (2012). Large-scale analysis of plasmid relationships through gene-sharing networks. *Molecular Biology and Evolution*, 29(4), 1225–1240. <https://doi.org/10.1093/molbev/msr292>
- Terpolilli, J. J., Hood, G. A., & Poole, P. S. (2012). What Determines the Efficiency of N₂-Fixing Rhizobium-Legume Symbioses? In *Advances in Microbial Physiology* (1st ed., Vol. 60). Elsevier Ltd. <https://doi.org/10.1016/B978-0-12-398264-3.00005-X>
- Thiele, I., & Palsson, B. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, 5(1), 93–121. <https://doi.org/10.1038/nprot.2009.203>
- Thies, J. E., Woome, P. L., & Singleton, P. W. (1995). Enrichment of Bradyrhizobium spp populations in soil due to cropping of the homologous host legume. *Soil Biology and Biochemistry*, 27(4–5), 633–636. [https://doi.org/10.1016/0038-0717\(95\)98643-3](https://doi.org/10.1016/0038-0717(95)98643-3)
- Thomas, C. M., & Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology*, 3(9), 711–721. <https://doi.org/10.1038/nrmicro1234>
- Thöny-Meyer, L. (1997). Biogenesis of respiratory cytochromes in bacteria. *Microbiology and Molecular Biology Reviews: MMBR*, 61(3), 337–376. <https://doi.org/10.1128/61.3.337-376.1997>
- Timmers, A. C. J., Soupene, E., Auriac, M. C., De Billy, F., Vasse, J., Boistard, P., & Truchet, G. (2000). Saprophytic intracellular rhizobia in alfalfa nodules. *Molecular Plant-Microbe Interactions*, 13(11),

- 1204–1213. <https://doi.org/10.1094/MPMI.2000.13.11.1204>
- Tiricz, H., Szücs, A., Farkas, A., Pap, B., Lima, R. M., Maróti, G., Kondorosi, É., & Kereszt, A. (2013). Antimicrobial nodule-specific cysteine-rich peptides induce membrane depolarization-associated changes in the transcriptome of *Sinorhizobium meliloti*. *Applied and Environmental Microbiology*, *79*(21), 6737–6746. <https://doi.org/10.1128/AEM.01791-13>
- Typas, A., Banzhaf, M., Gross, C. A., & Vollmer, W. (2012). From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature Reviews Microbiology*, *10*(2), 123–136. <https://doi.org/10.1038/nrmicro2677>
- Udvardi, M., & Poole, P. S. (2013). Transport and metabolism in legume-rhizobia symbioses. *Annual Review of Plant Biology*, *64*, 781–805. <https://doi.org/10.1146/annurev-arplant-050312-120235>
- Van De Velde, W., Guerra, J. C. P., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., Mergaert, P., Kondorosi, E., Holsters, M., & Goormachtig, S. (2006). Aging in legume symbiosis. A molecular view on nodule senescence in *Medicago truncatula*. *Plant Physiology*, *141*(2), 711–720. <https://doi.org/10.1104/pp.106.078691>
- Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A., Tiricz, H., Satiat-Jeunemaître, B., Alunni, B., Bourge, M., Kucho, K., Abe, M., Kereszt, A., Maróti, G., Uchiumi, T., Kondorosi, E., & Mergaert, P. (2010). Plant Peptides Govern Terminal Differentiation of Bacteria in Symbiosis. *Science*, *327*(February), 1122–1127. <https://doi.org/10.7551/mitpress/8876.003.0036>
- van Teeseling, M. C. F., de Pedro, M. A., & Cava, F. (2017). Determinants of bacterial morphology: From fundamentals to possibilities for antimicrobial targeting. *Frontiers in Microbiology*, *8*(JUL), 1–18. <https://doi.org/10.3389/fmicb.2017.01264>
- Vasse, J., De Billy, F., Camut, S., & Truchet, G. (1990). Correlation between ultrastructural differentiation of bacterioids and nitrogen fixation in alfalfa nodules. *Journal of Bacteriology*, *172*(8), 4295–4306. <https://doi.org/10.1128/jb.172.8.4295-4306.1990>
- Vences-Guzmán, M. A., Geiger, O., & Sohlenkamp, C. (2008). *Sinorhizobium meliloti* mutants deficient in phosphatidylserine decarboxylase accumulate phosphatidylserine and are strongly affected during symbiosis with Alfalfa. *Journal of Bacteriology*, *190*(20), 6846–6856. <https://doi.org/10.1128/JB.00610-08>
- Venkateshwaran, M., Jayaraman, D., Chabaud, M., Genre, A., Balloon, A. J., Maeda, J., Forshey, K., Os, D. Den, Kwiecien, N. W., Coon, J. J., Barker, D. G., & Ané, J.-M. (2015). A role for the mevalonate pathway in early plant symbiotic signaling. *Proceedings of the National Academy of Sciences*, *112*(38), E5378–E5378. <https://doi.org/10.1073/pnas.1516711112>
- Vinardell, J. M., Fedorova, E., Cebolla, A., Kevei, Z., Horvath, G., Kelemen, Z., Tarayre, S., Roudier, F., Mergaert, P., Kondorosi, A., & Kondorosi, E. (2003). Endoreduplication mediated by the anaphase-promoting complex activator CCS52A is required for symbiotic cell differentiation in *Medicago truncatula* nodules. *Plant Cell*, *15*(9), 2093–2105. <https://doi.org/10.1105/tpc.014373>
- Vincill, E. D., Szczyglowski, K., & Roberts, D. M. (2005). GmN70 and LjN70. Anion transporters of the symbiosome membrane of nodules with a transport preference for nitrate. *Plant Physiology*, *137*(4), 1435–1444. <https://doi.org/10.1104/pp.104.051953>
- Von Wintersdorff, C. J. H., Penders, J., Van Niekerk, J. M., Mills, N. D., Majumder, S., Van Alphen, L. B., Savelkoul, P. H. M., & Wolffs, P. F. G. (2016). Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Frontiers in Microbiology*, *7*(FEB), 1–10. <https://doi.org/10.3389/fmicb.2016.00173>

- Wang, D., Griffiths, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., Bisseling, T., & Long, S. (2010). A Nodule-Specific Protein Secretory Pathway Required for Nitrogen-Fixing Symbiosis. *Science*, *327*(February), 1126–1130.
- Wang, Q., Yang, S., Liu, J., Terecskei, K., Abrahám, E., Gombár, A., Domonkos, A., Szucs, A., Körmöczi, P., Wang, T., Fodor, L., Mao, L., Fei, Z., Kondorosi, E., Kaló, P., Kereszt, A., & Zhu, H. (2017). Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(26), 6854–6859. <https://doi.org/10.1073/pnas.1700715114>
- Wang, Z., Bie, P., Cheng, J., Lu, L., Cui, B., & Wu, Q. (2016). The ABC transporter YejABEF is required for resistance to antimicrobial peptides and the virulence of *Brucella melitensis*. *Scientific Reports*, *6*(July), 1–10. <https://doi.org/10.1038/srep31876>
- Wheatley, R. M., Ford, B. L., Li, L., Aroney, S. T. N., Knights, H. E., Ledermann, R., East, A. K., Ramachandran, V. K., & Poole, P. S. (2020). Lifestyle adaptations of *Rhizobium* from rhizosphere to symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(38), 23823–23834. <https://doi.org/10.1073/pnas.2009094117>
- Whiteley, M., Diggle, S. P., & Greenberg, E. P. (2017). Progress in and promise of bacterial quorum sensing research. *Nature*, *551*(7680), 313–320. <https://doi.org/10.1038/nature24624>
- Whitfield, G. B., Marmont, L. S., & Howell, P. L. (2015). Enzymatic modifications of exopolysaccharides enhance bacterial persistence. *Frontiers in Microbiology*, *6*(May), 1–21. <https://doi.org/10.3389/fmicb.2015.00471>
- Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: The unseen majority. *PNAS*, *95*, 6578–6583. [https://doi.org/10.1016/0006-8993\(80\)90730-1](https://doi.org/10.1016/0006-8993(80)90730-1)
- Wiles, T. J., Norton, J. P., Russell, C. W., Dalley, B. K., Fischer, K. F., & Mulvey, M. A. (2013). Combining Quantitative Genetic Footprinting and Trait Enrichment Analysis to Identify Fitness Determinants of a Bacterial Pathogen. *PLoS Genetics*, *9*(8). <https://doi.org/10.1371/journal.pgen.1003716>
- Wilson, A. C. C., & Duncan, R. P. (2015). Signatures of host/symbiont genome coevolution in insect nutritional endosymbioses. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(33), 10255–10261. <https://doi.org/10.1073/pnas.1423305112>
- Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews*, *51*(2), 221–271. <https://doi.org/10.1128/mmr.51.2.221-271.1987>
- Wojciechowski, M. F., Sanderson, M. J., Steele, K. P., & Liston, A. (2000). Molecular phylogeny of the “temperate herbaceous tribes” of papilionoid legumes: a supertree approach. *Advances in Legume ...*, *January*, 277–298. <http://www.public.asu.edu/~mfwojci/pdfs/WojoALS2000.pdf>
- Wozniak, R. A. F., & Waldor, M. K. (2010). Integrative and conjugative elements: Mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nature Reviews Microbiology*, *8*(8), 552–563. <https://doi.org/10.1038/nrmicro2382>
- Xia, Y., Suzuki, H., Borevitz, J., Blount, J., Guo, Z., Patel, K., Dixon, R. A., & Lamb, C. (2004). An extracellular aspartic protease functions in *Arabidopsis* disease resistance signaling. *EMBO Journal*, *23*(4), 980–988. <https://doi.org/10.1038/sj.emboj.7600086>
- Xie, F., Murray, J. D., Kim, J., Heckmann, A. B., Edwards, A., Oldroyd, G. E. D., & Downie, J. A. (2012). Legume pectate lyase required for root infection by rhizobia. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(2), 633–638. <https://doi.org/10.1073/pnas.1113992109>

- Xue, S., & Biondi, E. G. (2019). Coordination of symbiosis and cell cycle functions in *Sinorhizobium meliloti*. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1862(7), 691–696. <https://doi.org/10.1016/j.bbagr.2018.05.003>
- Yamaya-Ito, H., Shimoda, Y., Hakoyama, T., Sato, S., Kaneko, T., Hossain, M. S., Shibata, S., Kawaguchi, M., Hayashi, M., Kouchi, H., & Umehara, Y. (2018). Loss-of-function of ASPARTIC PEPTIDASE NODULE-INDUCED 1 (APN1) in *Lotus japonicus* restricts efficient nitrogen-fixing symbiosis with specific *Mesorhizobium loti* strains. *Plant Journal*, 93(1), 5–16. <https://doi.org/10.1111/tpj.13759>
- Yang, S., Wang, Q., Fedorova, E., Liu, J., Qin, Q., Zheng, Q., Price, P. A., Pan, H., Wang, D., Griffiths, J. S., Bisseling, T., & Zhu, H. (2017). Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proceedings of the National Academy of Sciences of the United States of America*, 114(26), 6848–6853. <https://doi.org/10.1073/pnas.1700460114>
- Yang, Y., Hu, X. P., & Ma, B. G. (2017). Construction and simulation of the *Bradyrhizobium diazoefficiens* USDA110 metabolic network: a comparison between free-living and symbiotic states. *Molecular BioSystems*, 13(3), 607–620. <https://doi.org/10.1039/C6MB00553E>
- Yano, K., Yoshida, S., Müller, 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., & Parniske, M. (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(51), 20540–20545. <https://doi.org/10.1073/pnas.0806858105>
- Young, J. M., Kuykendall, L. D., Martínez-Romero, E., Kerr, A., & Sawada, H. (2001). A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*. *International Journal of Systematic and Evolutionary Microbiology*, 51(1), 89–103. <https://doi.org/10.1099/00207713-51-1-89>
- Zakharov, S. D., Sharma, O., Zhalnina, M., Yamashita, E., & Cramer, W. A. (2012). Pathways of colicin import: Utilization of BtuB, OmpF porin and the TolC drug-export protein. *Biochemical Society Transactions*, 40(6), 1463–1468. <https://doi.org/10.1042/BST20120211>
- Zhao, L., Deng, Z., Yang, W., Cao, Y., Wang, E., & Wei, G. (2010). Diverse rhizobia associated with *Sophora alopecuroides* grown in different regions of Loess Plateau in China. *Systematic and Applied Microbiology*, 33(8), 468–477. <https://doi.org/10.1016/j.syapm.2010.08.004>
- Zipfel, C., & Oldroyd, G. E. D. (2017). Plant signalling in symbiosis and immunity. *Nature*, 543(7645), 328–336. <https://doi.org/10.1038/nature22009>

ANNEXES

ANNEX A: Generation and phenotyping of USDA110 mutants

For insertion mutagenesis, internal fragments (300-600 bp) of the target gene were PCR-amplified and cloned into the pVO155-nptII-Cefo-GFP vector (Okazaki et al., 2015). The resulting constructs were introduced into *B. diazoefficiens* USDA110 by conjugation and mutants were selected on 50 $\mu\text{g.mL}^{-1}$ kanamycin and 25 $\mu\text{g.mL}^{-1}$ cefotaxim plates. For the construction of deletion mutants, flanking regions of the gene were PCR amplified, fused by overlap extension PCR and cloned in plasmid pFLY (a modified pNPTS129 suicide plasmid harboring a YFP gene). The resulting plasmids were introduced into *B. diazoefficiens* USDA110 and mutants formed after a double recombination event, were obtained by counterselection on 10% sucrose medium and further screening for the loss of plasmid-derived antibiotic resistance and fluorescence. The genotype of all mutant strains was verified by PCR analysis. The phenotype of the mutants was analyzed by inoculating *G. max* and *A. afraspera* plants as indicated above. Fourteen dpi plants were collected, and their leaf and nodule color and number were visually inspected. Nitrogenase activity was determined by the acetylene reduction assay as previously described (Barrière et al., 2017).

ANNEX B: Tn-Seq, a Transposon Insertion Sequencing technique

Transposon insertion sequencing (TIS) comprise a set of functional genomics techniques, which combine transposon insertional mutagenesis with massive parallel sequencing of the transposon insertion sites. They aim to identify bacterial genes contributing to a function of interest. Four different techniques were originally developed: HITS (High throughput Insertion Track by deep Sequencing), INSeq (Insertion Sequencing), TraDIS (Transposon-Directed Insertion site Sequencing) and Tn-seq (Transposon Sequencing; **Fig. annex B**). Each has its

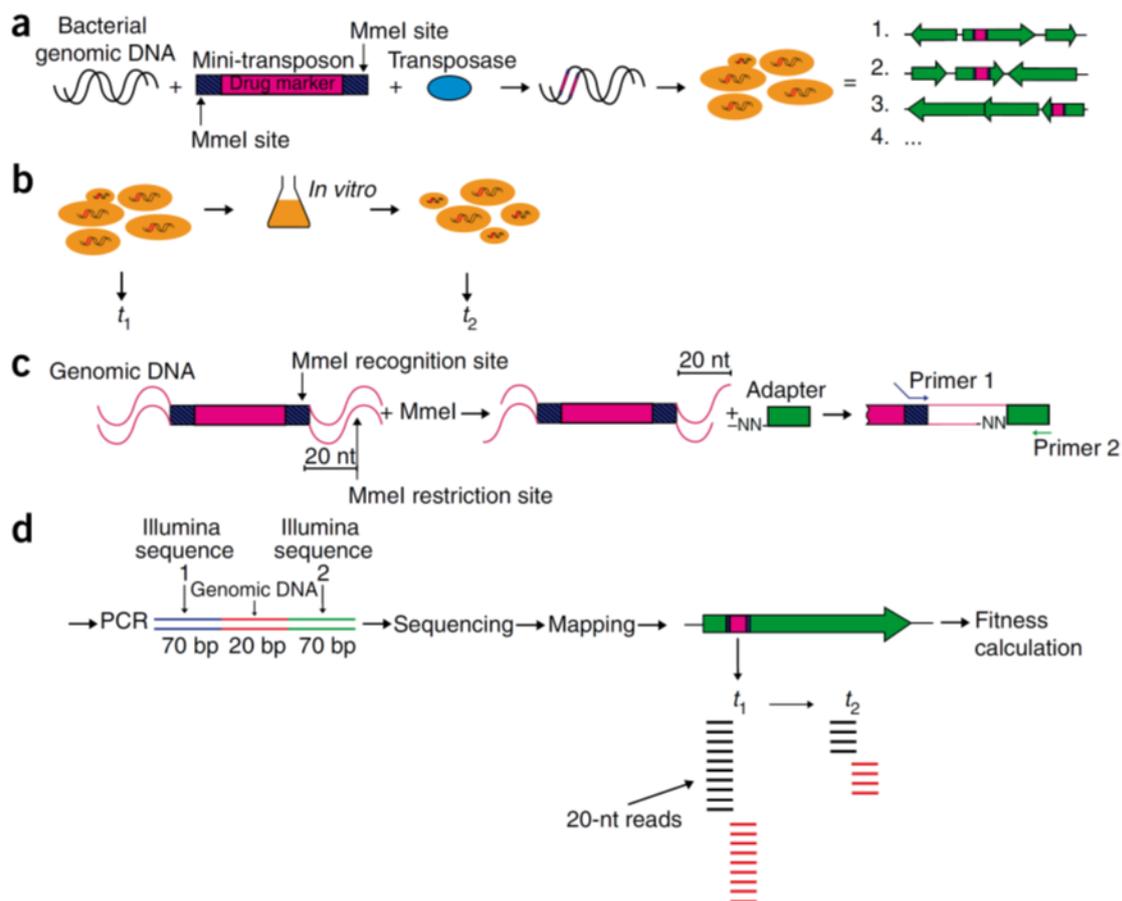


Figure annex B | Schematic representation of the Tn-Seq procedure. a | A library of transposon interruption mutants is built through *in vitro* transposition. The transposon carries a drug resistance marker and is flanked with MmeI recognition sites. After transformation of bacteria with the transposed DNA, the library is obtained in which every bacterium contains a single copy of the transposon. **b** | DNA is isolated from the raw library (t_1) while a pool of mutant from the library is grown on the test condition. DNA of this selected library subset is extracted as well (t_2). **c** | DNA from t_1 and t_2 is digested with MmeI, which cuts 20bp aside its recognition site and adapters are ligated to the digestion products. **d** | The ligation products are amplified by PCR, yielding a 160-bp sequence comprising the 20 bp of bacterial-specific DNA. This fragment is flanked with Illumina sequences, necessary for the sequencing. Libraries t_1 and t_2 are subsequently sequenced and the 20-bp reads are mapped to the genome allowing counting the number of insertions for each TA site and from which the fitness of each mutant can be deduced. Extracted from van Opijnen et al., 2009

technical specificities, such as transposon usage or the length of the sequenced insertion site. All of them are based on the assembly of a saturated transposon insertion library.

In Tn-Seq the transposons used are of the Himar I mariner type which specifically insert at TA sites. The library is then grown on a test condition, and the flanking regions of transposons are sequenced. To do so, we take advantage of a MmeI recognition site located on the extremities of the transposon. This restriction enzyme is of the type III as it cuts 20 base pairs away from the recognition site. Using this feature, we are able to isolate 20bp of the bacterial chromosomes on each transposon border. Adding Illumina adapters by ligation and PCR amplification allows to generate the fragments ready for sequencing. This *en masse* sequencing of the library allows to assess the frequency of each insertion mutant which reflects their fitness. Strong decrease in the frequency of insertion at a TA site means the mutant is unable to grow and that the site is essential in the test condition.

ANNEX C: Metabolic modelling of ORS285

Metabolic models (MM) allow to reconstruct the theoretical metabolite synthesis pathways and to simulate the growth of a bacterium. A MM is basically a list of reactions and metabolites which are fed to a mathematic prediction program. The software will try to grow the modeled bacterium *in silico*. Different techniques are available to do so, such as the FBA (Flux Balance Analysis) method. An FBA is a stoichiometry-based solving algorithm, which uses a branch of mathematics called optimization (**Fig. annex C – 1**). In the simplest cases, the role of optimization is to model, analyze and solve problems, which consists in maximizing or minimizing a function over a given set. In a MM, the rate from each reaction is considered a parameter of a complex function, which takes as input metabolite concentration in the media and outputs biomass production. These programs will usually tweak each reaction rates in order to maximize the growth of the bacteria (deduced from biomass production) with the input metabolites. To run the ORS285 model, we used the CoBRA toolbox (Constraint Based Reconstruction and Analysis) (Heirendt et al., 2019).

We built the model by following described procedures in the literature (DiCenzo et al., 2016; Thiele & Palsson, 2010; **Fig. annex C - 2**), we first, using annotation files of the organism, we built an initial draft model thanks to the Kbase online software. This tool assesses the reactions present base in one organism based on gene annotation and proposes a list of metabolites required or produced in these reactions. It also fills the gaps in the proposed pathways based on a provided growth media composition, so the model is able to grow in this media. However, as an automated process, this model is far from reflecting the real metabolism of our bacteria. A manual curation is therefore needed. Having the chance to have MM of relatively closely related bacteria (*S. meliloti* 1021 and *B. diazoefficiens* USDA110), we refined the model using a BBH (Bidirectional Best Hits) BLAST (Basic Local Alignment Search Tool) approach. Reactions from the two cited MM performed by genes having orthologs in ORS285 were added to our MM. This yields an expanded draft MM of ORS285 for which directionality of newly added reactions were predicted (using Kbase) and redundant reactions removed. After these steps, the final draft model was obtained, ready for the real manual curation.

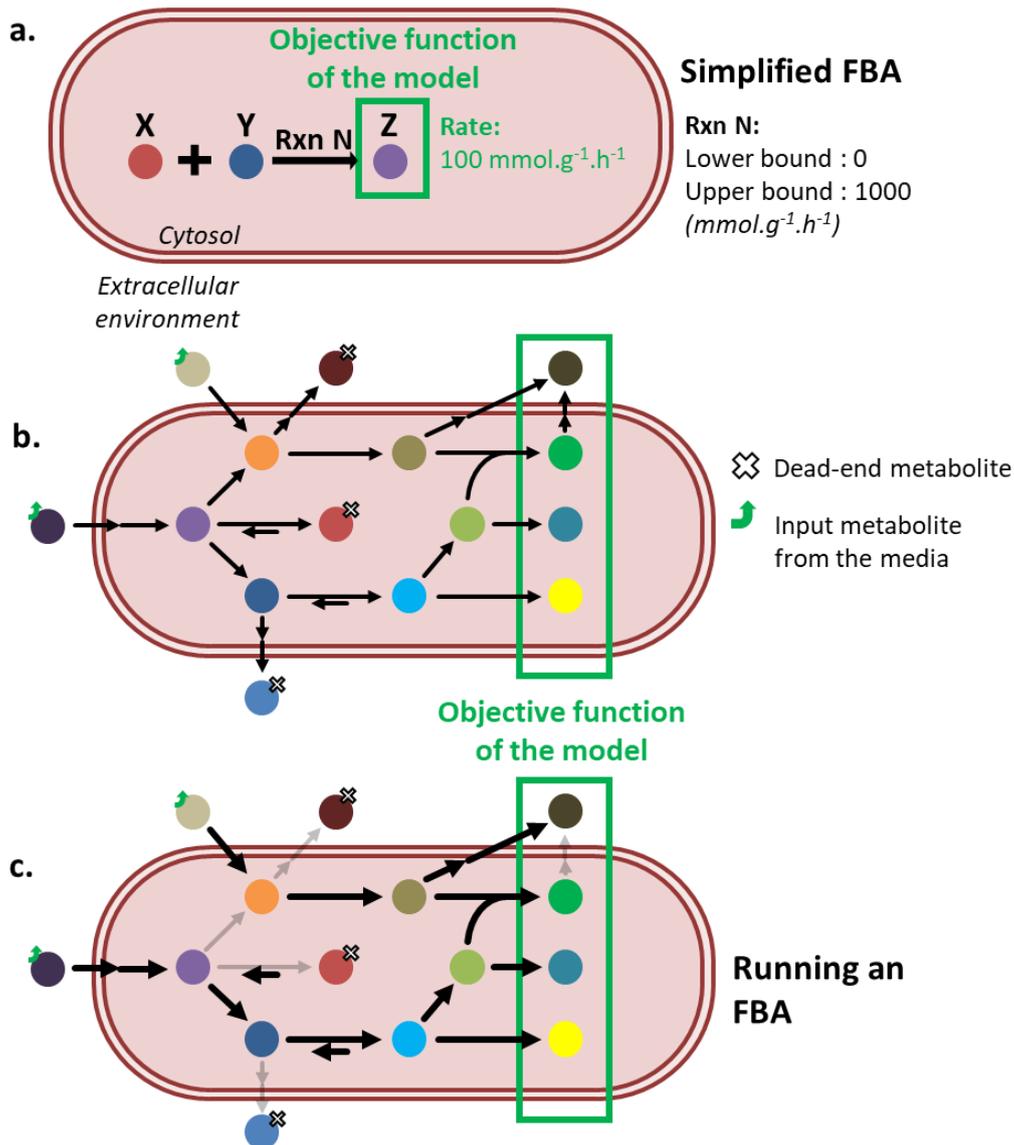


Figure annex C - 1 | Simplified representation of an FBA. a | Bacterial model simplified to one reaction. Each reaction is defined by the list of reactants, products, directionality and flux rate boundaries ($\text{mmol.g}^{-1}.\text{h}^{-1}$). In this example reaction, N has defined flux rate limits and is unidirectional, processing metabolites X and Y into Z. In order to infer the growth of the model, an objective function must be specified. This function defines a minimal production rate for the set of metabolites considered to be necessary to generate two bacteria from one. If an FBA is run with such a model, the software will tweak the value of the reaction N flux rate, within the fixed boundaries to maximize the growth of the model. It will search for the reaction NN flux rate values that maximizes the production of metabolites in the objective function, here Z. IN this example, the FBA will likely set the reaction N flux rate to $1000 \text{ mmol.g}^{-1}.\text{h}^{-1}$, therefore allowing the model to grow. **b |** Complexification of the model presented in a. In this representation are added the metabolites given as input and simulating the growth medium composition. As the model processes only flux rates, we add them to the media at a certain rate. Therefore, simulating the growth of a model is closer to bio-fermenter than culture in an Erlenmeyer. Dead-end metabolites are also displayed in this representation, which represent metabolites that are produced but not processed (the reverse also exist). Whereas the objective function is more complex in this case, it has the same goal than in the previous representation. **c |** Running the FBA in this complexified model will have the same outcomes. The software tweaks the reaction rates for each reactions trying to maximize the growth of the model. Dead-end metabolites are often excluded from the analysis as being sink metabolites.

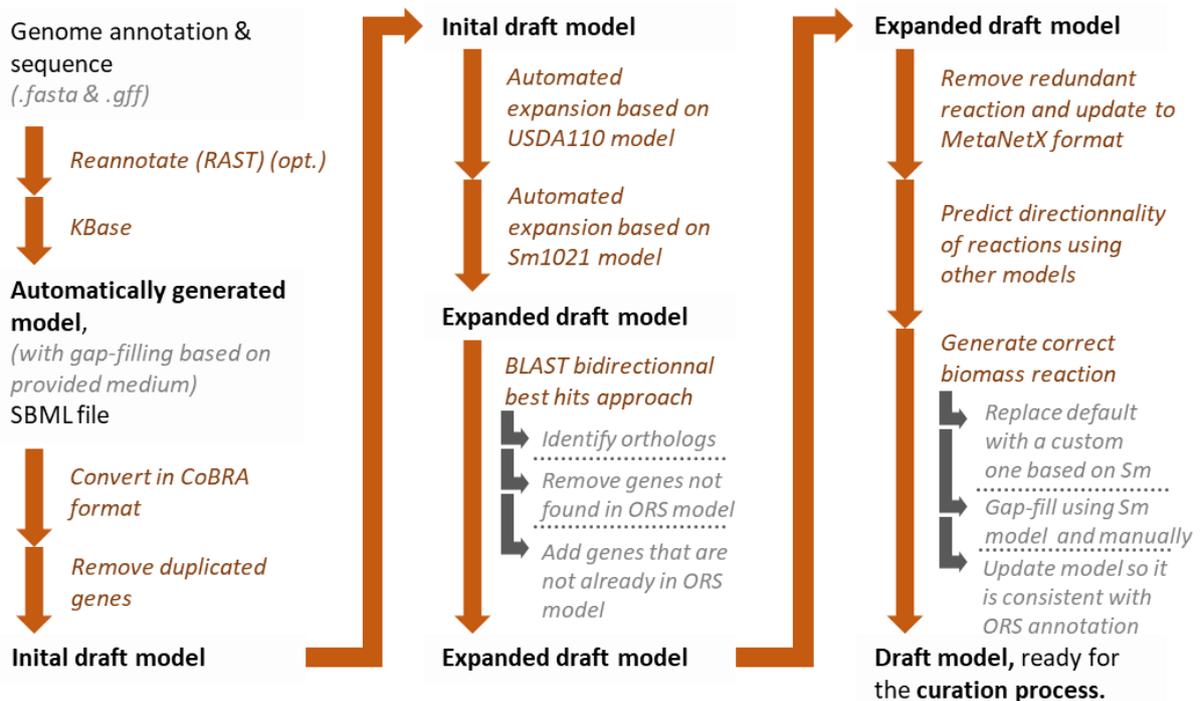


Figure annex C - 2 | Pipeline used to build ORS285 draft model. The optional automatic reannotation of the genome was performed. The annotation format was not compatible with KBase automated model generation. However, it introduced a lot of incorrect annotation in the model, which required to be manually corrected. SBML: System Biology Markup Language (file format of the model); Sm: *S. meliloti*.

The manual curation allows to get as close as possible to the actual metabolism of the bacteria. This process is based on experimental data. First, functions that were known as specific in the bacterium were added to the model. This comprises a specific sugar bradyrhizose (Busset et al., 2017; Silipo et al., 2011), hopanoids (Kulkarni et al., 2015) and hopanoids linked to the LPS (Silipo et al., 2014). Next, biochemistry data such as enzymatic activity would have been a good way to keep refining the model. However, limited data is available on *Bradyrhizobium* species in this field. As an alternative, we decided to use two approaches to refine the model.

One of them is the use of Biolog-equivalent growth assays various carbon sources. A Biolog assay is usually a method to quickly identify a microbe based on his metabolic functions. By growing a bacterium in different substrates, it is possible to assess which carbon sources the organism can metabolize. Through this approach we compared the experimental results to the growth of the model in the same conditions. Each difference between the *in silico* and experimental results means the model is not accurate enough in the related metabolic pathways. If the model grows whereas the bacterium does not mean some reactions might be

deleted (as could be related metabolites), or their directionality modified. On the contrary, if the model cannot grow when the bacterium, *ti* could mean that some reactions are lacking or that their directionality is wrong.

Using this method, we assessed both *in vitro* and *in silico* growth with 190 compounds and nitrate as the nitrogen source. A total of 58 compounds were supporting growth *in vitro*, however the model predicted growth with only 21. Over 132 compounds not supporting growth, the model predicted an absence of growth 119. After manually correcting the model, 41 compounds could support growth and 132 could not. The 17 compounds for which growth is different between ORS285 and its model would not be solved through manual curation. Independently verification of balance and charge was carried out. It allowed to gain the ability to grow with D-Glucuronic acid, such as in the Biolog assays. However, growth with D-mannitol malonic acid were lost. Adding back a D-mannitol dehydrogenase reaction and changing directionality of an ATP synthase reaction respectively solved these growth problems. At this point, the model displays incorrect growth only for 16 compounds.

Besides the Biolog assays method, we decided to use Tn-seq as a way to curate the association of genes to reactions. Indeed, in the final draft models, and thanks to the biology assays, though reactions begin to be representing ORS285 behavior accurately enough, many reactions are still associated with multiple genes. To resolve this problem a similar strategy than for Biolog assays was carried out. The idea was to generate Tn-seq dataset for the growth in minimal media supplied with various carbon and nitrogen sources, which would allow us to assess the gene essentiality for the growth with each carbon and nitrogen source. Similarly, we generated *in silico* single mutants for each gene in the model and assessed their essentiality for the growth in equivalent media. We generated the ORS285 transposon mutant library by conjugating the pSAM_Ec plasmid (Wiles et al., 2013) which carries the mariner-like transposon into the target strain. Afterwards, mutants were selected using two different media: a rich medium (YM, used in chapter A of the results) and a minimal medium (BNM-Modified). BNM-Modified is derived from the aquaponic solution used for the study of *Aeschynomene* species and ORS285 interactions using succinate and ammonium nitrate as carbon and nitrogen sources. Growing conjugants is already a selection *per se*, and thus depletes some mutants in the library. I wanted to know whether we could reduce this bias. To do so, I mixed both libraries

with a 50:50 ratio (cell:cell) and used it along with the other libraries to compare their essential genes. Sadly, the minimal media proved to be unusable for the Tn-Seq experiment as there was no correlation in the read counts for each gene between our replicates and ORS285 grew incredibly slow in this medium. Investigation the growth of ORS285 in different growth media is undergoing. Sadly, due to a lack of time and to the sanitary situation, I could not go further in this project, which I hope will continue to progress in the frame of a collaboration with George diCenzo (now settled in Queen's University in Canada).

Titre : Étude de fonctions bactériennes impliquées dans la différenciation bactérienne lors de la symbiose rhizobium-légumineuse

Mots clés : Symbiose Rhizobium-Légumineuse, Bactéroïdes, Différenciation terminale, Peptides NCR, Enveloppe bactérienne, Multiomiques

Résumé : La symbiose rhizobium-légumineuse est une interaction étroite entre plante et bactérie. Au cours de cette symbiose, la bactérie est hébergée par la plante au sein d'organes symbiotiques où elle fixe l'azote atmosphérique. Les espèces de légumineuses du groupe des IRLC et des Dalbergioïdes peuvent contrôler les rhizobia symbiotiques et induire un processus de différenciation particulier grâce à la production massive de peptides riches en cystéines (NCR) spécifiques aux nodosités. *In vitro*, les peptides NCR cationiques ont des activités de perméabilisation membranaire sur de nombreuses bactéries. La manière dont les rhizobia s'adaptent pour résister à ce stress intense reste encore aujourd'hui mal comprise. Deux axes de recherche principaux ont été menés au cours de cette thèse, tous deux liés à la compréhension de la réponse des bactéries à la différenciation terminale imposée par les peptides NCR.

D'un côté, nous avons analysé le rôle de certaines fonctions bactériennes dans la résistance aux NCR et au cours de l'interaction entre *Medicago truncatula* et *Sinorhizobium meliloti*. Dans ce travail, nous avons principalement évalué les fonctions membranaires telles que la synthèse du LPS, le système de réponse aux stress de l'enveloppe et des fonctions d'importation. Nous avons trouvé de nouvelles fonctions qui pourraient être impliquées dans la résistance à la NCR et la différenciation terminale des bactéroïdes.

De l'autre côté, nous avons mené une approche multi-omique couplée à des techniques de biologie cellulaire pour caractériser l'interaction mal adaptée entre *Bradyrhizobium diazoefficiens* USDA110 et *Aeschynomene afraspera*. Nous avons découvert de nouvelles particularités dans cette interaction avec notamment une différenciation inhabituelle.

Title: Study of terminal bacteroid differentiation features during the legume-rhizobium symbiosis.

Keywords: Legume-rhizobium symbiosis, Bacteroids, Terminal differentiation, NCR peptides, Bacterial envelope, Multiomics

Abstract: The legume-rhizobia symbiosis is a close interaction between plants and bacteria. During this symbiosis, bacteria are hosted by the plants in symbiotic organs called nodules. There, the symbionts fix atmospheric nitrogen for the plants. Legume species from IRLC and Dalbergioids can control symbiotic rhizobia and mediate a particular differentiation process through the massive production of nodule-specific cysteine rich (NCR) peptides. *In vitro*, cationic NCR peptides have membrane-permeabilizing activities on many bacteria. How rhizobia adapt to resist this intense stress remains poorly understood. Two main research axes were driven during this thesis, both linked to the understanding of how bacteria react to terminal differentiation imposed by NCR peptides.

On one side, we analyzed the role of bacterial functions in NCR resistance and during interaction between *Medicago truncatula* and *Sinorhizobium meliloti*. In this work, we mainly assessed membrane functions such as LPS synthesis, Envelope Stress Response, and import functions. We found novel functions that could be involved in NCR resistance and terminal bacteroid differentiation.

On the other side, we conducted a multi-omics approach coupled with cell-biology techniques to characterize the ill-adapted interaction between *Bradyrhizobium diazoefficiens* USDA110 and *Aeschynomene afraspera*. We discovered new features in this interaction with an unusual differentiation.