Analyse comparative des mécanismes de différenciation des bactéroïdes au cours des symbioses Bradyrhizobium Aeschynomene

Florian Lamouche

▶ **To cite this version:**

Florian Lamouche. Analyse comparative des mécanismes de différenciation des bactéroïdes au cours des symbioses Bradyrhizobium Aeschynomene. Biologie végétale. Université Paris Saclay (COmUE), 2019. Français. NNT : 2019SACLS036 . tel-02426228

HAL Id: tel-02426228
https://theses.hal.science/tel-02426228

Submitted on 2 Jan 2020

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.

université PARIS-SACLAY

NNT : 2019SACLS036

Thèse de doctorat

UNIVERSITÉ PARIS SUD
Comprendre le monde, construire l'avenir

Analyse comparative des mécanismes de différenciation des bactéroïdes au cours des symbioses *Bradyrhizobium-Aeschynomene*

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°567
Sciences du végétal : du gène à l'écosystème
Spécialité de doctorat : Biologie

Thèse présentée et soutenue publiquement à Gif-sur-Yvette le 01/02/2019 par

Florian Lamouche

Composition du Jury :

Valérie Hocher Chargée de recherche, IRD (LSTM)	Rapporteur
Emanuele Biondi Directeur de recherche, CNRS (LCB)	Rapporteur
Claude Bruand Directeur de recherche, INRA (LIPM)	Examinateur
Frédérique van Gijsegem Chargée de recherche, INRA (IEES)	Examinateur
Soufian Ouchane Directeur de recherche, CNRS (I2BC)	Examinateur Président du jury
Benoit Alunni Maitre de Conférences, Université Paris-Sud (I2BC)	Directeur de thèse

Directeurs de thèse :

Benoit Alunni Maitre de Conférences, Université Paris-Sud (I2BC)	Directeur de thèse
Yves Dessaux Directeur de recherche, CNRS (I2BC)	Co-directeur de thèse

Résumé

Titre : Analyse comparative des mécanismes de différenciation des bactéroïdes au cours des symbioses *Aeschynomene-Bradyrhizobium*

Mots clés : Symbiose fixatrice d'azote, bactéroïdes, polyploïdie, transcriptomique, protéomique, métabolomique, efficacité symbiotique

Résumé : En cas de carence azotée, les légumineuses sont capables de mettre en place une symbiose avec des bactéries du sol fixatrices d'azote appelées rhizobia. Cette symbiose a lieu dans un organe appelé nodosité où les bactéries sont endocytées et appelées bactéroïdes. Certains clades de légumineuses imposent un processus de différenciation à leurs bactéroïdes qui agrandissent considérablement et deviennent polyploïdes, menant à des morphotypes bactériens allongés ou sphériques. Au cours de cette thèse, j'ai étudié la différenciation des bactéroïdes de *Bradyrhizobium* spp. en association avec *Aeschynomene* spp.. Les bactéroïdes de ces plantes présentent des degrés de différenciation distincts qui dépendent de l'espèce hôte. Mes données suggèrent que les bactéroïdes les plus différenciés sont aussi les plus efficaces. J'ai cherché à savoir quels facteurs procaryotes pourraient être impliqués dans les adaptations des bactéroïdes au processus de différenciation et à leurs divers hôtes, le tout en lien avec cette différence d'efficacité symbiotique au travers d'approches globales sans *a priori* de type -omiques. Les conditions considérées sont des bactéroïdes de différents morphotypes et des cultures libres de référence. Les fonctions activées en conditions symbiotiques ont été identifiées, ainsi que les gènes spécifiques d'un hôte donné. Des analyses fonctionnelles des gènes d'intérêt ont également été menées. Les mutants bactériens n'ont toutefois pas présenté de phénotype symbiotique drastique, montrant ainsi l'existence de réseaux de gènes complexes menant à la résilience des génomes de rhizobia.

Title : Comparative analysis of bacteroid differentiation mechanisms in *Aeschynomene-Bradyrhizobium* symbioses

Keywords : Nitrogen-fixing symbiosis, bacteroids, polyploidy, transcriptomics, proteomics, metabolomics, symbiotic efficiency

Abstract: In case of nitrogen starvation, legume plants establish a symbiotic interaction with nitrogen-fixing soil bacteria called rhizobia. This interaction takes place in nodules where the symbionts are internalized and become bacteroids. Some legume clades also impose a differentiation process onto the bacteroids which become enlarged and polyploid, leading to elongated or spherical morphotypes. During my PhD work, I have studied bacteroid differentiation of *Bradyrhizobium* spp. in association with *Aeschynomene* spp.. These bacteroids display distinct differentiation levels depending on the plant host, and my analyses suggest that the most differentiated ones are also the most efficient. I investigated the bacterial factors potentially involved in the adaptations to differentiation and host-specificity, and related to the higher efficiency of the most differentiated bacteroids using global-omics approaches without *a priori*. The analyzed conditions were bacteroids of distinct morphotypes and free-living reference cultures. Activated functions under symbiotic conditions were identified, as well as host-specific ones. Functional analyses were performed on genes of interest. However, the bacterial mutants did not display drastic symbiotic phenotypes, showing the existence of complex gene networks leading to high resilience of rhizobial genomes.

Remerciements

« C'était sur une grande route, je marchais là depuis des jours,
Voire des semaines ou des mois, je marchais là depuis toujours.
Une route pleine de virages, des trajectoires qui dévient,
Un chemin un peu bizarre, un peu tordu comme la vie »

Dire que cette thèse a été un virage d'envergure relève presque de l'euphémisme. Alors que cette aventure touche à sa fin, je tiens à remercier ici toutes les personnes qui ont eu un rôle déterminant dans la réalisation de mon travail de thèse, tant en termes de compétences scientifiques que de soutien humain.

En premier lieu je tiens à remercier mon directeur de thèse Benoit Alunni, pour le temps qu'il a pris dans mon encadrement, les relecture d'articles et de ce manuscrit, ainsi que pour tout ce que j'ai pu apprendre auprès de lui. Mes pensées vont ensuite à Yves Dessaux et Peter Mergaert, qui ont co-dirigé cette thèse, que cela soit officiellement ou non. Merci pour ces discussions constructives et fructueuses tout au long de ce travail.

Je tiens ensuite à remercier tous les membres de mon jury – Valérie Hocher, Emanuele Biondi, Claude Bruand, Frédérique van Gijsegem et Soufian Ouchane – d'avoir accepté d'évaluer mon travail.

Je remercie également les Marie-Laure Martin Magniette, Marie Anne Barny et Eric Giraud pour avoir fait partie de mon comité de thèse. Leurs avis éclairés ont eu un rôle déterminant dans les directions prises au cours de ce travail.

Je tiens également à remercier tous membres permanents de l'équipe « Interactions Plantes-Bactéries » : Tania Timtchenko, Denis Faure, Jeremy Cigna, Catherine Grandclément et Raynald Cossard. Je vous remercie pour votre aide et les discussions que nous avons eus pendant ma thèse. Un grand merci à Corine Foucault pour sa bienveillance et son aide précieuse.

Je remercie tous les membres non-permanents actuels ou anciens de l'équipe « Interactions Plantes-Bactéries » et du bâtiment 34 qui ont croisé mon chemin pendant ce long périple qu'a été ma thèse. Merci Pauline Blin, Nicolas Busset, Grégory Deicsics, Almudena Gonzalez Mula, Ibtissem Guefrachi, Slimane Khayi, Joy Lachat, Quentin Nicoud, Tsubasa Ohbayashi, Said Oulghazi, Olivier Pierre, Kévin Robic, Jonathan Torel-Salat et Marta Torres Bejar. Je me dois d'adresser une attention particulière à Quentin Barrière, qui fut un collègue et un véritable ami tout au long de cette épopée.

Les résultats de cette thèse sont le fruit d'un travail collaboratif avec d'autres équipes de recherche. Merci à Thierry Balliau, Mélisande Blain-Nicolas et Michel Zivy de la plate-forme protéomique PAPPSO de la Ferme du Moulon. Mes remerciements vont à Christophe Salon et Christian Jeudy de l'INRA Dijon, pour notre collaboration sur la fluxomique. Je remercie également Bertrand Gakière, Florence Guérard et Françoise Gilard de la plateforme de Métabolomique de l'IPS2. Je pense aussi à Attila Kereszt, István Nagy et Attila Szücs du BRC de Szeged pour avoir généré les données de RNA-seq.

Je remercie tout le personnel de la plateforme « Imagif », et particulièrement Mickaël Bourge, Nicolas Valentin, Romain Le Bars, Laetitia Besse et Cynthia Gillet pour leur assistance technique et leur bienveillance.

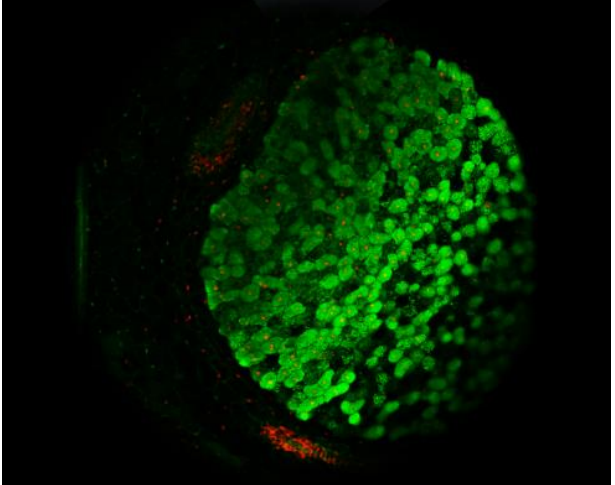
Je remercie Eric Giraud (à nouveau), Djamel Gully, Joël Fardoux et Albin Teulet pour m'avoir également chaleureusement accueilli au sein du LSTM et transmis leurs connaissances.

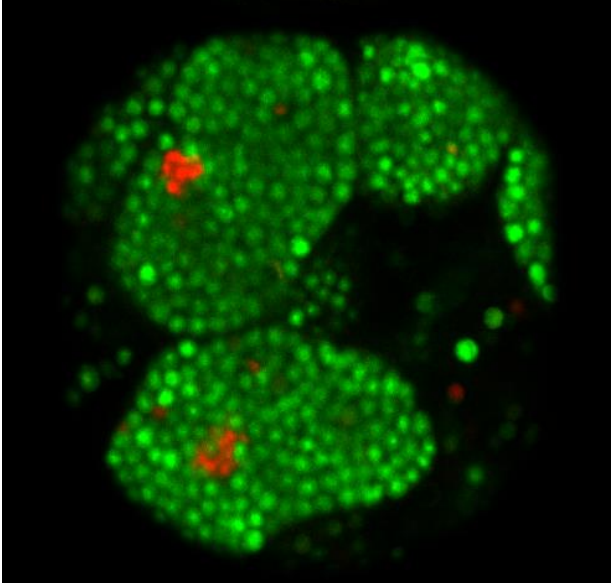
J’exprime ici ma reconnaissance envers le Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation pour le financement de ma bourse de thèse.

Je n’oublie pas toutes les personnes qui petit à petit, m’ont aidé à construire ce projet de devenir docteur en sciences. Je pense à l’ensemble de mes enseignants, toutes disciplines confondues, qui m’ont donné goût à l’apprentissage au point de faire de moi leur futur collègue.

Enfin, je suis éternellement reconnaissant envers ma famille et mes ami·es, toutes ces personnes qui comptent pour moi, que j’ai rencontré tout au long de ce long parcours, et avec qui j’espère encore faire un bout de chemin.

« J'ai rencontré quelques peines, j'ai rencontré beaucoup de joie
C'est parfois une question de chance, souvent une histoire de choix
Je ne suis pas au bout de mes surprises, là-dessus il n’y a aucun doute
Et tous les jours je continue d'apprendre les codes de ma route »





Sommaire

Sommaire

Sommaire 2

Table des illustrations 5

Partie I : Introduction 10

1. La fixation biologique de l'azote 10

1.1. Les enjeux de la fixation de l'azote 10

1.1.1. L'azote, brique essentielle du vivant 10

1.1.2. La nutrition azotée des végétaux terrestres 12

1.1.3. La fertilisation des sols en agriculture et la fixation artificielle de l'azote 13

1.1.4. Conséquences négatives de l'utilisation des engrais azotés et enjeux 14

1.2. Le complexe enzymatique clé de la fixation biologique : la nitrogénase 17

1.3. La diversité des organismes diazotrophes 19

1.4. Les symbioses fixatrices d'azote 23

1.4.1. Interactions symbiotiques et origines des Eucaryotes 23

1.4.2. Exemples de symbioses fixatrices d'azote avec des métazoaires 23

1.4.3. Symbioses fixatrices d'azote entre cyanobactéries et Embryophytes 24

1.4.4. Symbioses fixatrices d'azote entre Angiospermes et bactéries rhizosphériques ou endophytes 25

1.4.5. Les interactions endosymbiotiques entre Protéobactéries et Angiospermes 26

2. Les symbioses rhizobium-légumineuses 28

2.1. Généralités sur les légumineuses 28

2.1.1. Diversité et écologie des légumineuses 28

2.1.2. Importance agronomique des légumineuses 30

2.2. La rhizosphère et ses habitants 32

2.2.1. La rhizosphère, un lieu d'interactions entre plantes et microorganismes 32

2.2.2. La structure génomique des rhizobia et des gènes symbiotiques 33

2.2.3. La diversité des rhizobia et la variabilité de leur spectre d'hôte 36

2.3. Le dialogue moléculaire entre les légumineuses et les rhizobia 37

2.3.1. La libération de flavonoïdes et leur perception par les rhizobia 37
2.3.2. La sécrétion de facteurs Nod initie l'interaction entre les partenaires 38
2.3.3. La voie de signalisation commune des endosymbioses 40
2.3.4. La formation des cordons d'infection et morphogénèse des nodosités 41
2.3.5. La différenciation des cellules symbiotiques 45
2.3.6. L'expression des gènes bactériens impliqués dans la fixation d'azote 46
2.3.7. Les échanges nutritionnels hôte-symbionte au niveau des symbiosomes 47
3. Les processus de différenciation cellulaire des cellules symbiotiques et des bactéroïdes 49
3.1. La diversité des morphotypes de bactéroïdes 49
3.3. La différenciation des bactéroïdes 50
3.3.1. Observations cytologiques 50
3.3.2. Mécanismes moléculaires responsables de la différenciation : les peptides NCR 52
3.3.3. Les protéines bactériennes impliquées dans la différenciation des bactéroïdes 56
4. La symbiose *Bradyrhizobium-Aeschynomene* 60
4.1. Diversité des *Aeschynomene* et établissement de groupes d'inoculation croisés 60
4.2. Les processus de nodulation indépendante des facteurs Nod chez les *Aeschynomene* 62
4.2.1. Découverte d'un processus symbiotique Nod-indépendant 62
4.2.2. Les processus Nod-indépendants et T3SS-dépendant 63
4.2.3. La signalisation Nod-indépendante active certains acteurs de la voie de signalisation commune des endosymbioses 65
4.3. Infection et organogénèse des nodosités chez les *Aeschynomene* 66
4.4. La différenciation des bactéroïdes d'*Aeschynomene* 69
4.4.1. Diversité des morphotypes des bactéroïdes et différenciation chez les symbiontes d'*Aeschynomene* 69
4.4.2. Implication de peptides NCR-like d'*Aeschynomene* dans le processus de différenciation 70
4.4.3. Identification de BclA, un ABC-transporteur bactérien requis pour la différenciation des bactéroïdes 71
5. L'apport des données -omiques et notion(s) d'efficacité symbiotique 73
5.1. Comparaison des différentes approches -omiques 73

5.2. Apport des approches -omiques dans la compréhension des mécanismes symbiotiques entre rhizobia et légumineuses 75

5.3. L'efficacité symbiotique dans les associations Rhizobium-Légumineuses 77

5.3.1. Des questions de définition, d'échelle et de point de vue 77

5.3.2. Valeur sélective bactérienne et différenciation des bactéroïdes 79

5.3.3. Différenciation et efficacité symbiotique du point de vue végétal 80

5.4. Problématiques et objectifs du projet de thèse 81

Partie II : Résultats 86

1. Symbiotic efficiency of spherical and elongated bacteroids in the *Aeschynomene-Bradyrhizobium* symbiosis 86

2. Transcriptomic dissection of *Bradyrhizobium* sp. strain ORS285 in symbiosis with *Aeschynomene* spp. inducing different bacteroid morphotypes with contrasted symbiotic efficiency 101

3. From intracellular bacteria to differentiated bacteroids: transcriptome and metabolome analysis in *Aeschynomene* nodules using the *Bradyrhizobium* sp. ORS285 Δ*bclA* mutant 153

4. Atypical terminal differentiation of *Bradyrhizobium diazoefficiens* USDA110 bacteroid in *Aeschynomene afraspera* analyzed through multi-omic approaches 184

Partie III : Discussion 222

Références bibliographiques 250

Annexes 268

1. Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium* strains 268

2. Integrated roles of BclA and DD- carboxypeptidase 1 in *Bradyrhizobium* differentiation within NCR-producing and NCR-lacking root nodules 286

3. The biotroph *Agrobacterium tumefaciens* thrives in tumors by exploiting a wide spectrum of plant host metabolites 304

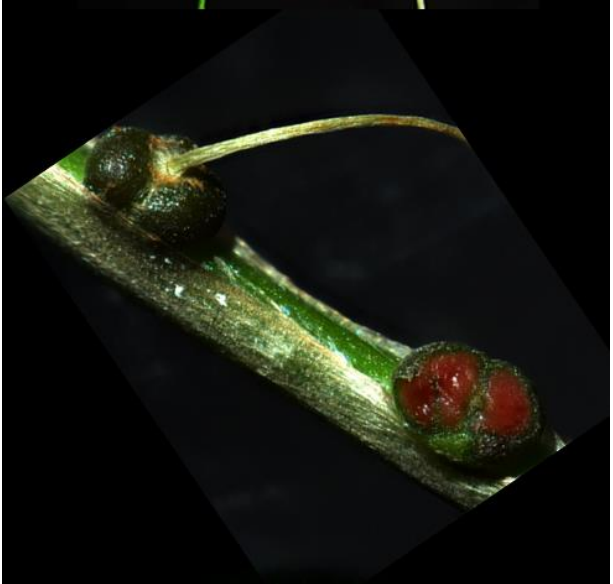
4. Comparative cytology, physiology and transcriptomics of free-living and symbiotic *Burkholderia* reveal features for colonization of its insect host 323

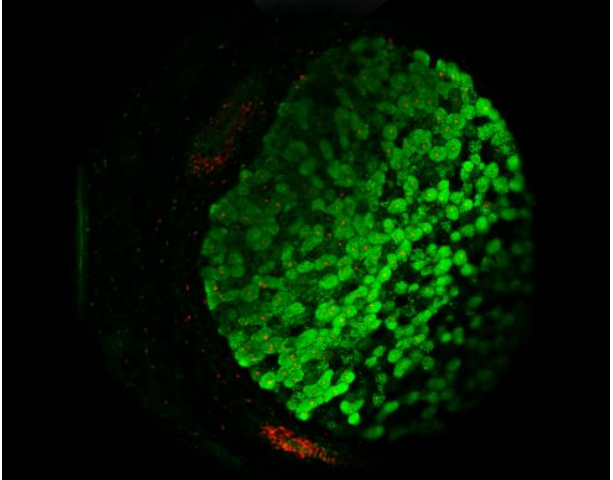
Table des illustrations

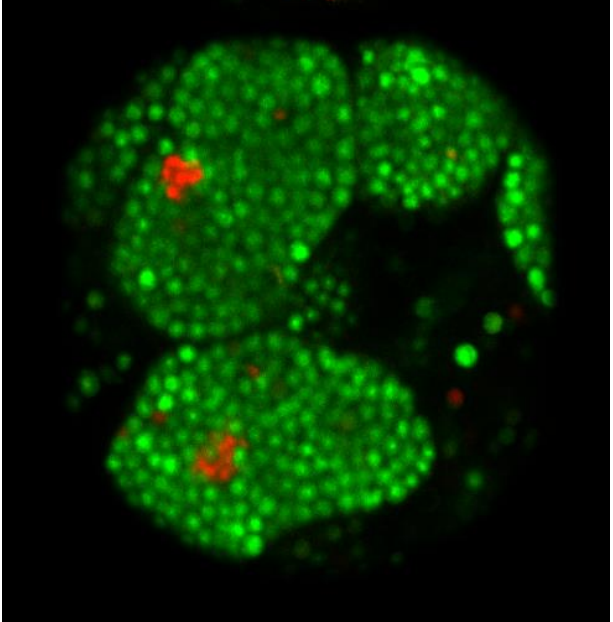
Figure 1: Les transformations majeures dans le cycle de l'azote 11
Figure 2 : Formes, origines et devenir de l'azote réactif dans le sol 12
Table 1 : Différents types d'engrais azotés et leurs contenus en azote 14
Figure 3: Flux d'azote naturel et lié aux perturbations anthropiques aux niveaux continental et océanique 15
Figure 4: Effet négatifs de l'utilisation des engrais azotés et phosphatés sur les écosystèmes 15
Figure 5: Estimations futures des impacts négatifs liés à l'utilisation des engrais azotés dans le monde 16
Figure 6 : Organisation génomique et fonctionnelle du complexe nitrogénase 18
Figure 7: Représentation schématique des trois domaines du vivant, mettant l'accent sur les lignées Procaryotes capables de fixation d'azote. 20
Figure 8: Les cyanobactéries fixatrices d'azote. 21
Figure 9: Les actinomycètes du genre *Frankia*. 22
Figure 10: Les symbioses fixatrices d'azote entre Embryophytes et cyanobactéries. 25
Figure 11: Représentation des associations symbiotiques possibles entre bactéries réalisant la fixation biologique de l'azote et Embryophytes 27
Figure 12: Nouvelle classification des *Fabaceae* basée sur la phylogénie des séquences *matK* 29
Figure 13: Evolution des rendements mondiaux en céréales et en légumineuses à graines de 1961 à 2013 31
Figure 14: Attachement et colonisation des racines de légumineuses par les rhizobia 33
Figure 15: Organisation génomique des rhizobia 34
Figure 16: Arbre phylogénétique de séquences d'ADNr 16S provenant d'α-, β- et γ-protéobactéries 35
Figure 17: Phylogénie basée sur l'ADNr 16S des Protéobactéries montrant la distribution des rhizobia 37
Figure 18: L'induction des gènes de nodulation par les flavonoïdes 38
Figure 19: Structure du facteur Nod de *Ensifer fredii* NGR234 et protéines impliquées dans sa synthèse 39
Figure 20: Schéma simplifié de la réponse à la perception des facteurs Nod chez *Medicago truncatula* 40
Figure 21: Infection des poils racinaires de *Medicago truncatula* par *Ensifer meliloti* 42
Figure 22: Différents modes d'infection des racines par les rhizobia parallèles à l'organogénèse de nodosités fixatrices d'azote 43
Figure 23: Structure des nodosités racinaires déterminées et indéterminées 44
Figure 24: Les cellules symbiotiques des nodosités de *M. truncatula* subissent des cycles d'endoréplication requise pour une symbiose fonctionnelle, et cette transition est permise par la protéine Ccs52A 45
Figure 25: Comparaison des cascades de régulation de l'expression des gènes de fixation d'azote des rhizobia entre *E. meliloti* 1021 (A) et *B. diazoefficiens* USDA110 (B) 47
Figure 26: Schéma récapitulatif des échanges nutritionnels intervenants au niveau du symbiosome entre bactéroïde et cellule symbiotique 48
Figure 27: Diversité des morphotypes de bactéroïdes chez les légumineuses 49
Figure 28: Phylogénie de 40 espèces de *Papilionoideae* résumant le type de nodosité et l'occurrence du processus de différenciation des bactéroïdes 50
Figure 29: Le processus de différenciation des bactéroïdes dans les nodosités d'IRLC (Inverted Repeat Lacking Clade) 51

Table 2: Répartition spatiale des différents types de bactéroïdes le long d'une nodosité indéterminée et leurs caractéristiques 51
Figure 30: Comparaison de la taille et du niveau de ploïdie des cellules d'*Ensifer meliloti* 1021 en vie libre ou en symbiose avec *M. truncatula* 52
Figure 31: Similitudes dans la structure primaire entre les peptides NCR et des peptides antimicrobiens de type défensine 53
Figure 32: NCR001 colocalise avec les symbiosomes chez le génotype sauvage de *M. truncatula*, mais pas chez le mutant *dnf1* 54
Figure 33: Les peptides NCR provoquent des caractéristiques de différenciation terminale en vie libre 55
Figure 34: La différenciation des bactéroïdes fait intervenir le régulateur « maitre » CtrA 57
Figure 35: Phénotype symbiotique des nodosités infectées par le mutant bactérien *ΔbacA* d'*E. meliloti* 58
Figure 36: Mécanismes moléculaires de la mise en place lors du processus de différenciation des bactéroïdes chez *E. meliloti* 59
Figure 37: Comparaisons phénotypiques et histologiques entre les nodosités de soja (*Glycine max*, gauche) et *Aeschynomene indica* (droite) 61
Figure 38: Groupes d'inoculation croisés entre bradyrhizobia symbiotiques et leurs hôtes *Aeschynomene* compatibles 62
Figure 39: Phénotypes de nodulation des souches ORS278, ORS285 et ORS285*ΔnodB* avec *A. afraspera* et *A. sensitiva*. 63
Figure 40: Rôle du système de sécrétion de type III dans l'initiation de l'interaction symbiotique entre *Bradyrhizobium elkanii* USDA61 et ses plantes hôtes 64
Figure 41: Modèle simplifié de la voie de signalisation symbiotique entre rhizobia et légumineuses 65
Figure 43: Infection et organogénèse des nodosités d'*Aeschynomene indica* inoculées par la souche *Bradyrhizobium* sp. ORS285 67
Figure 44: Infection et organogénèse des nodosités d'*Aeschynomene afraspera* inoculées par la souche *Bradyrhizobium* sp. ORS285 68
Figure 44: La différenciation cellulaire des bactéroïdes de *Bradyrhizobium* sp. ORS285 en symbiose avec *A. afraspera* et *A. indica* 70
Figure 45: La différenciation des bactéroïdes d'*Aeschynomene* est médiée par des peptides NCR-like 71
Figure 46: BclA est requis dans le processus de différenciation des bactéroïdes d'ORS285 avec les hôtes *A. afraspera* et *A. indica* 72
Figure 47: Représentation schématique des trois niveaux de caractérisation des systèmes biologiques par des approches -omiques 74
Figure 48: Résumé des systèmes expérimentaux fréquemment utilisés lors d'approches –omiques appliquées aux bactéries symbiotiques de légumineuses 76
Figure 49: Cycle de vie de l'interaction symbiotique entre rhizobia et une légumineuse annuelle, et paramètres à quantifier pour mesurer l'efficacité et le bénéfice de l'interaction pour les deux partenaires. 79
Figure 50: Comparaison de l'efficacité symbiotique de bactéroïdes différenciés et non différenciés 81
Figure 51: Systèmes expérimentaux et comparaisons effectuées au cours des analyses transcriptomiques (T) et protéomiques (P) 83

Figure 52: Modèles intégrant la diversité et l'intensité des processus de différenciation des bactéroïdes existant chez les IRLC et les Dalbergioïdes, basée sur une échelle de niveaux de ploïdie. 233
Figure 53: Comparaisons morphologiques de culture de la souche CzR2 dans différents milieux de culture. ... 234
Figure 54: Boucles de régulation médiées par CtrA et démontrées via diverses méthodes 242
Figure 55: Stratégies actuelles visant à transférer la fixation biologique de l'azote chez des plantes de grande culture autres hors-légumineuses. 246







Introduction

Partie I : Introduction

1. La fixation biologique de l'azote

1.1. Les enjeux de la fixation de l'azote

1.1.1. L'azote, brique essentielle du vivant

Quatrième élément le plus abondant du vivant, l'azote entre dans la composition de nombreuses biomolécules. Tout d'abord, c'est un élément majeur des nucléotides, constitués d'un pentose phosphate et d'une base azotée. Ces nucléotides sont les monomères des acides nucléiques qui, polymérisés, sont les formes de stockage (ADN) et d'expression (ARN) de l'information génétique des organismes vivants. L'azote est également présent dans les protéines constituées d'acides aminés possédant une ou deux fonctions amines ($-NH_2$). Un dernier exemple de molécules azotées jouant un rôle majeur sont les chlorophylles, qui opèrent la conversion de l'énergie solaire en énergie chimique. Elles sont à la base de la photosynthèse qui permet l'assimilation du carbone minéral (sous forme de CO_2) et donc sa transformation en carbone organique (sous forme de sucres). Les organismes chlorophylliens autotrophes au carbone constituent les producteurs primaires, c'est-à-dire la base des réseaux trophiques des écosystèmes terrestres et marins. L'azote est également le principal composant de l'atmosphère terrestre actuelle, sous forme de diazote gazeux N_2 (79% de l'air ambiant).

Cependant, malgré son abondance, il n'est pas directement assimilable par la plupart des êtres vivants. La quantité totale d'azote d'origine biologique dans les océans et sur les continents est estimée à 10^6 tonnes, ce qui représente un millionième du réservoir du N_2 atmosphérique (Ward, 2012). L'ensemble des formes d'azote assimilables par la majorité des organismes représente lui 0,001% de l'azote de la biosphère (Newton, 1998). Cet élément est de fait souvent limitant dans la productivité des écosystèmes et notamment pour les producteurs primaires. Le diazote atmosphérique devient assimilable pour les organismes vivants via un processus appelé « fixation d'azote » qui peut être biologique ou non. L'azote peut ainsi être rendu assimilable par les éclairs qui, en libérant jusqu'à $4.\ 10^{10}$ Watts, favorisent la réaction de transformation du diazote en oxydes d'azote NO_x (Borucki et Chameides, 1984), qui peuvent être directement assimilés. La fixation biologique de l'azote n'est quant à elle effectuée que par

certains organismes Procaryotes dits diazotrophes. Ce processus est crucial pour l'ensemble des êtres vivants car il participe à la régulation naturelle de la quantité de biomasse terrestre.

L'importance de la fixation biologique est d'autant plus grande qu'elle est la première étape d'entrée de l'azote dans un cycle global de modifications chimiques (Figure 1). Les bactéries fixatrices d'azote réduisent le diazote en ammonium NH_4^+ assimilable par d'autres organismes (végétaux, champignons, bactéries). Cet ammonium peut ensuite être oxydé sous forme de nitrite (NO_2^-) puis de nitrate (NO_3^-) par d'autres bactéries. Ce processus appelé nitrification, est effectué par les bactéries nitrifiantes. Par exemple, dans le sol, les bactéries du genre *Nitrosomonas* et *Nitrobacter* oxydent respectivement le NH_4^+ en NO_2^- et le NO_2^- en NO_3^-. L'azote sous forme assimilable (NH_4^+ et NO_3^-) est ensuite absorbé par les producteurs primaires, qui le convertissent en acides aminés. Cela permet leur croissance et donc une augmentation de leur biomasse, dont la consommation constitue la base des réseaux trophiques. La mort des organismes susmentionnés fait retourner l'azote dans l'océan et/ou le sol où il peut à nouveau être dégradé en ammonium par des bactéries saprophytes. La dernière étape du cycle est le retour sous la forme de N_2 gazeux appelée dénitrification. Elle est provoquée par des bactéries dénitrifiantes et anammox qui sont des bactéries anaérobies tirant leur énergie de la réduction des oxydes d'azote (Strous *et al.*, 1999). Ainsi la grande majorité des interconversions observées au sein du cycle biogéochimique de l'azote est effectuée par des organismes vivants Procaryotes.

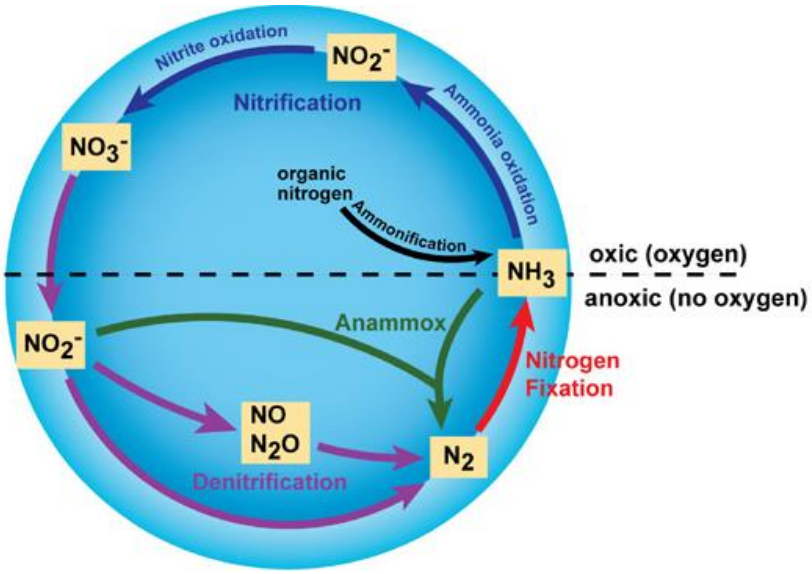


Figure 1: Les transformations majeures dans le cycle de l'azote (d'après Bernhard, 2010).

1.1.2. La nutrition azotée des végétaux terrestres

Les interconversions d'azote d'origine Procaryote sont nécessaires à la nutrition des plantes terrestres. La majorité de l'azote du sol est de nature organique et se trouve dans les molécules complexes de l'humus, qui sont possiblement converties en NH_4^+ par des microorganismes (Procaryotes et champignons) du sol selon le processus appelé minéralisation (Miller et Cramer, 2004). Ces organismes se nourrissent essentiellement de nitrate (NO_3^-), d'ammonium (NH_4^+) et d'acides aminés libres présents dans le sol selon des proportions dépendantes des conditions physico-chimiques inhérentes aux sols (Figure 2). De manière générale, les forêts matures ou celles de la toundra arctique sont établies sur des sols acides et/ou réducteurs, menant les plantes à prélever préférentiellement l'ammonium ou des acides aminés libres. Les plantes adaptées aux sols alcalins semblent quant à elles prélever plus facilement des nitrates (Maathuis, 2009). L'équilibre crée par ce cycle biogéochimique limitait, avant la période industrielle, la quantité d'azote assimilable disponible pour la biosphère.

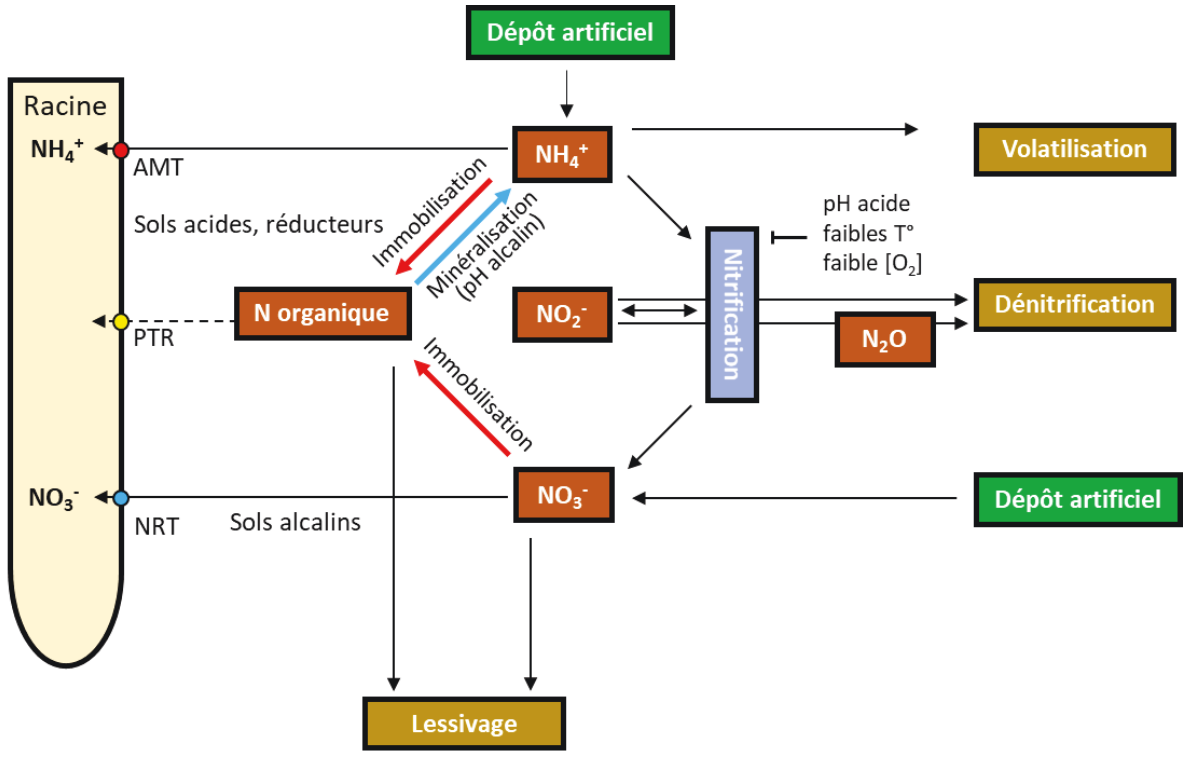


Figure 2 : Formes, origines et devenir de l'azote réactif dans le sol. Le processus d'immobilisation de l'azote sous forme organique est majoritairement dû à la décomposition de la matière organique d'origine animale et végétale sous forme d'humus. Le processus de minéralisation est permis par des Procaryotes et des champignons du sol. Les dépôts artificiels proviennent des différents engrais azotés (voir Table 1). AMT (*Ammonium Transporters*), PTR (*Peptide transporters*) et NRT (*Nitrate transporters*) sont les différentes familles de transporteurs racinaires de NH_4^+, acides aminés libres et NO_3^-, respectivement (d'après Millet et Cramer 2004 ; Maathuis, 2009).

1.1.3. La fertilisation des sols en agriculture et la fixation artificielle de l'azote

Depuis l'apparition de l'agriculture au Néolithique (Demoule, 2010), on peut distinguer trois types d'actions de l'homme sur le cycle de l'azote. Tout d'abord, des méthodes de fertilisation empirique des sols ont vu le jour, via le recouvrement des terres cultivées par les excréments des animaux domestiqués, l'agriculture sur brûlis et l'utilisation de plantes de la famille de légumineuses. Ces techniques n'ont sans doute pas drastiquement modifié les flux d'azote au sein du cycle. Cependant, l'équilibre élémentaire des sols a été grandement modifié après plusieurs siècles d'agriculture. En effet, les sols sont des systèmes ouverts qui s'appauvrissaient en azote car la quantité d'intrants azotés déposés était inférieure à la production sortante. Les rendements, notamment en blé d'hiver en France, ne cessaient ainsi de diminuer au XVII^{e} et XIX^{e} siècles, alors que la population ne cessait de croître (Boulaine, 1995a ;1995b). De nombreux progrès scientifiques, notamment en chimie et en agronomie, ainsi que des découvertes géologiques ont mené à une seconde phase, remobilisant de l'azote enfoui dans la biosphère et donc dans le cycle de l'azote. L'exploitation et l'importation d'engrais naturels tels que le guano péruvien ou les nitrates de sodium du Chili ont par exemple permis d'appliquer des intrants nécessaires en azote (mais également en phosphore et potassium), restaurant considérablement la fertilité des sols des pays Européens (Boulaine, 1995).

La troisième phase s'est opérée avec l'avènement des engrais chimiques de synthèse, basés sur la fixation artificielle du diazote atmosphérique par le procédé de Haber-Bosch inventé au début du XX^{e} siècle (Smil, 2001). Cette invention est considérée comme l'une des plus importantes de l'humanité. Elle constitue un des piliers de la révolution verte qui a contribué à la croissance démographique, menant à un quadruplement de la population mondiale en un siècle (Smil, 2001). C'est un procédé coûteux en énergie qui synthétise de l'ammoniac (NH_3) par hydrogénation du diazote (N_2). Pour optimiser les rendements d'un point de vue cinétique, cette réaction nécessite de fortes élévations de température et de pression (~500°C et 15-25 MPa). La réaction chimique s'effectue selon l'équation bilan : $N_{2(g)} + 3\ H_{2(g)} \leftrightarrows 2\ NH_{3(g)}$. Ce procédé a été utilisé dès 1913 en Allemagne et signe le point de départ de la chimie de l'azote à l'échelle industrielle, qui permet aujourd'hui une production mondiale d'ammoniac de 130 mégatonnes par an dont 80% sont destinés à la production d'engrais azotés (Smil, 2001). Les

engrais azotés de synthèse les plus utilisés ainsi que leurs contenus en azote, sont présentés dans la Table 1 (Cao *et al.*, 2017).

Type d'engrais	Formule chimique	Contenu en azote (%)	Forme d'azote
Ammoniac anhydre	NH_3	82	NH_4^+
Ammoniaque	NH_3	22	NH_4^+
Nitrate d'ammonium	NH_4NO_3	34	50% NH_4^+ et 50% NO_3^-
Sulfate d'ammonium	$(NH_4)_2SO_4$	21	NH_4^+
Solutions azotées	$CO(NH_2)_2$, NH_4NO_3	30	75% NH_4^+ et 25% NO_3^-
Nitrate de sodium	$NaNO_3$	16	NO_3^-
Urée	$CO(NH_2)_2$	46	NH_4^+
Nitrate de calcium	$Ca(NO_3)_2$	17	NO_3^-
Diammonium phosphate	$(NH_4)_2HPO_4$	18	NH_4^+
Monoammonium phosphate	$NH_4H_2PO_4$	11	NH_4^+
Ammonium Phosphates	$(NH_4)_3PO_4$, etc	15	NH_4^+

Table 1 : Différents types d'engrais azotés et leurs contenus en azote (d'après Cao *et al.*, 2017).

1.1.4. Conséquences négatives de l'utilisation des engrais azotés et enjeux

L'utilisation massive d'engrais, bien que garantissant une meilleure sécurité alimentaire pour l'Homme, a exercé et exerce toujours actuellement des effets désastreux sur l'environnement (Horrigan *et al.*, 2002). Cette fixation anthropique a altéré le cycle de l'azote (Figure 3, Gruber & Galloway, 2008) et a mené à de nombreuses complications. Il est admis que 2 à 10% de l'azote fixé artificiellement se retrouve réellement dans l'alimentation humaine (Fields, 2004). Les engrais constitués de NH_4^+ induisent une volatilisation rapide sous forme de NH_3 gazeux, tandis qu'une grande partie des engrais nitratés n'est pas absorbée par les cultures, mais est lessivée dans les cours d'eau (Cao *et al.*, 2017). Ceci induit une très forte teneur en nitrates de l'eau, ce qui provoque la pollution des nappes phréatiques et favorise la croissance d'algues et de plantes aquatiques pouvant mener à l'eutrophisation en eau douce et dans les estuaires (Figure 4 ; Rastogi *et al.*, 2015). La non-régulation d'une telle pollution au fil des années conduit à une diminution de la biodiversité et à la destruction d'écosystèmes. C'est notamment le cas du golfe du Mexique où se déversent les eaux riches en nitrates du Mississipi. Les marées vertes d'algues et leur décomposition privent les eaux plus profondes d'oxygène, créant une zone morte où la vie aquatique aérobie est devenue impossible (Fields, 2004). Par ailleurs, l'azote retenu dans le sol peut subir la dénitrification et libérer à la fois du diazote (N_2)

et du protoxyde d'azote (N_2O) dans l'atmosphère.

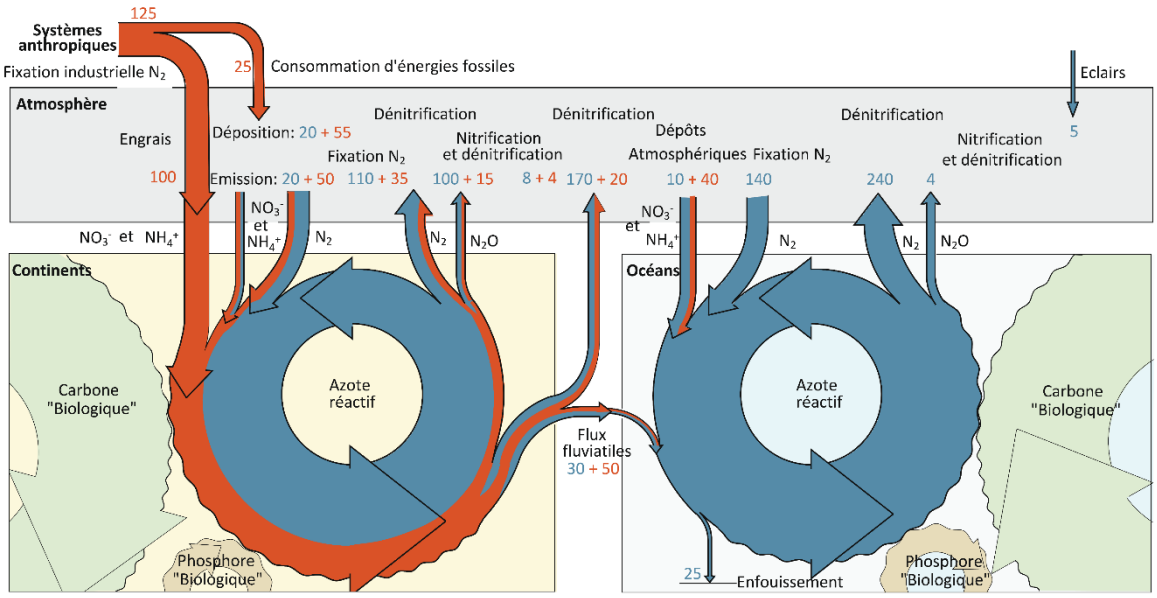


Figure 3: Flux d'azote naturel et lié aux perturbations anthropiques aux niveaux continental et océanique. Les flux naturels et artificiels (en mégatonnes) sont respectivement indiqués en bleu et rouge (d'après Gruber & Galloway, 2008).

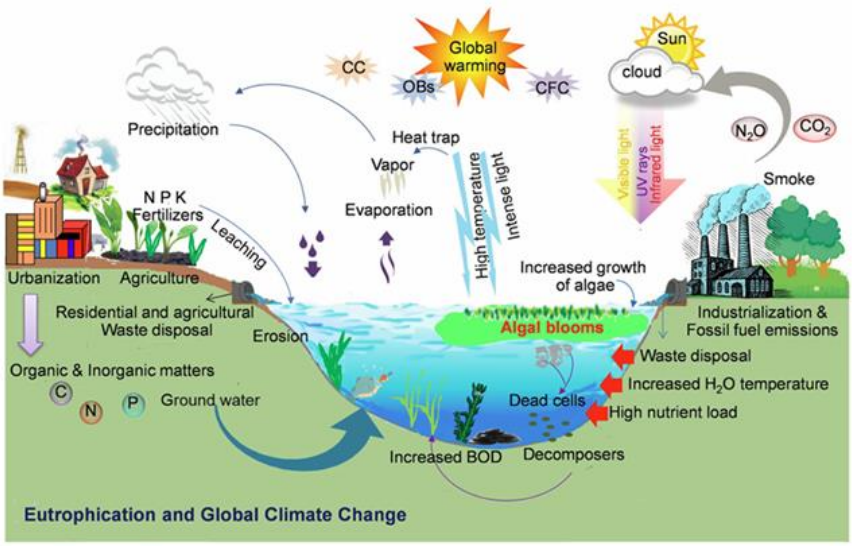


Figure 4: Effet négatifs de l'utilisation des engrais azotés et phosphatés sur les écosystèmes. Schéma illustrant les facteurs clés induisant les marées vertes d'algues (d'après Rastogi *et al.*, 2015).

Certains auteurs estiment que 2% et 2,5% de l'azote provenant respectivement du fumier et des engrais azotés ont été transformés en N_2O entre 1860 et 2005 (Davidson, 2009). Ces pollutions diverses liées aux engrais azotés ont des conséquences sur la santé humaine. Par exemple, certaines études considèrent que le réchauffement climatique induit par les modifications anthropiques du cycle de l'azote favorise l'expansion de maladies infectieuses

comme la malaria, en étendant la niche écologique de leurs vecteurs (Erisman *et al.*, 2013).

La volatilisation des engrais azotés formant du NH_3 et des oxydes d'azotes NO_x augmente la formation d'ozone troposphérique qui est associée à plus de 21 000 morts prématurées et 14 000 hospitalisations dues à des problèmes respiratoires chaque année dans les pays membres de l'Union Européenne (Amann *et al.*, 2008 ; Erisman *et al.*, 2013). De surcroît, il devient de plus en plus clair que le réchauffement climatique, en partie causé par l'utilisation de ces engrais, a des répercussions négatives sur les rendements agricoles. Par exemple, une augmentation de 1°C de la température sur Terre induirait une diminution des rendements des cultures de blé de 5 à 15% (Zhao *et al.*, 2017, Figure 5A). Il y a eu depuis des études décrivant et analysant les effets provoqués par l'excès des flux anthropiques d'azote et de phosphore sur l'environnement. De nombreux scientifiques ont tenté de modéliser leur impact présent et futur sur la biosphère afin de proposer une ligne de conduite basée sur des limites non viables à ne pas dépasser sur le long terme (Steffen *et al.*, 2015, Figure 5B).

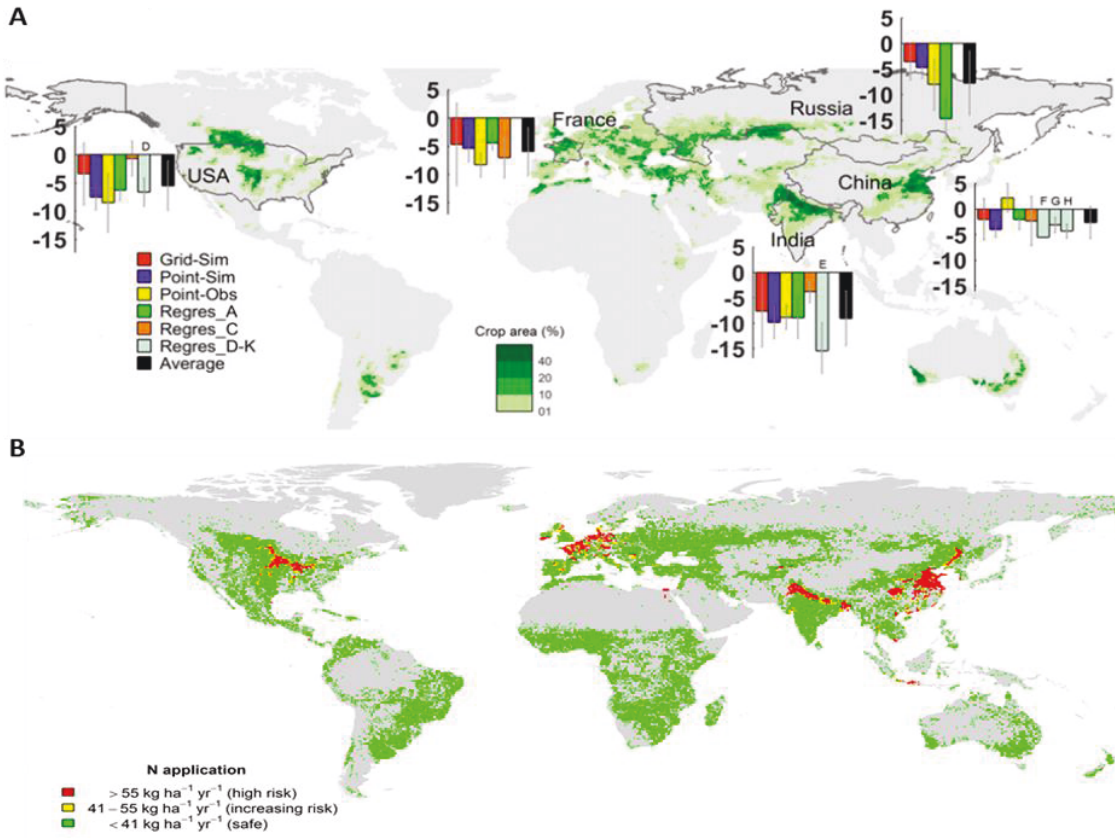


Figure 5: Estimations futures des impacts négatifs liés à l'utilisation des engrais azotés dans le monde. (A) Résultats de différentes modélisations mesurant les impacts négatifs du réchauffement climatique sur les rendements agricoles de blé. Chaque couleur est un modèle estimant le pourcentage de perte de rendements en cas de la hausse de température globale de 1°C dans 5 régions du monde (d'après Zhao *et al.*, 2017). (B) Carte mondiale des zones à risques concernant les flux d'azote lié à l'utilisation d'intrants azotés (d'après Steffen *et al.*, 2015).

Malheureusement, la situation est assez dramatique concernant les perspectives à long terme et l'humanité doit faire face à un défi double : maintenir une production agricole suffisante pour nourrir une population mondiale en constante augmentation, tout en limitant au maximum l'utilisation de ces engrais dont les effets délétères sont de mieux en mieux caractérisés. La réussite de ce défi passera sans doute par des changements politiques courageux et drastiques concernant nos modes de consommation et de production agricole. Néanmoins, succéder à la puissance et l'efficacité à court terme de l'industrie chimique ne se fera pas sans progrès scientifiques qui permettront un maintien des rendements tout en évitant les perturbations anthropiques décrites plus haut. Au moment de son discours de réception du prix Nobel de Chimie en 1918, Fritz Haber a ainsi déclaré que son invention n'était sans doute pas le point final du travail à mener sur l'azote, et qu'il faudrait sans doute regarder du côté de sa fixation biologique (Smil, 2001). C'est en effet une piste intéressante qui génère un regain d'intérêt au sein du monde agricole. Les innovations associées à l'étude de ce processus autoriseraient une meilleure gestion du cycle de l'azote en termes de durabilité tout en garantissant une sécurité alimentaire aux populations humaines toujours en expansion (Alunni et Mergaert, 2017).

1.2. Le complexe enzymatique clé de la fixation biologique : la nitrogénase

La fixation biologique de l'azote n'est effectuée que par des organismes Procaryotes grâce à un complexe enzymatique : la nitrogénase (Seefeldt *et al.*, 2009). Elle catalyse la réaction chimique suivante : $N_2 + 8H^+ + 16MgATP + 8^{e-} \rightarrow 2NH_3 + H_2 + 16MgADP + 16Pi$. Le complexe nitrogénase le mieux caractérisé est celui de la bactérie *Klebsiella oxytoca* (anciennement *K. pneumoniae*). La nitrogénase est constituée principalement de deux métalloprotéines. La première, un homodimère appelé dinitrogénase réductase ou encore Fe-protéine, est codée par le gène *nifH*. Cette protéine présente une masse moléculaire de 64kDa et possède deux sites de fixation d'ATP. La deuxième protéine, dite Mo-Fe, est un hétérotétramère nommé dinitrogénase 2α2β codée par les gènes *nifD* (sous-unité α) et *nifK* (sous-unité β) de 250kDa. Au sein de la Fe-protéine se situe un cluster [4Fe4S], qui transfère un électron à la protéine Molybdène-Fer (Mo-Fe). Dans la protéine-MoFe, pour chaque unité αβ, un cluster P reliant deux complexes Fer-Soufre [8Fe7S] par un atome S central est présent (Seefeldt *et al.*, 2009). Il est considéré comme impliqué dans le transfert des électrons entre les deux métalloprotéines de la nitrogénase, réceptionnant ainsi l'électron du cluster [4Fe4S] de la

protéine-Fe. Ce transfert d'électron dissocie ensuite les deux métalloprotéines jusqu'à recyclage de la Fe-protéine. Dans la protéine-MoFe, l'accepteur final des électrons est le cofacteur [7FeMo9S :C]-homocitrate, appelé FeMo-Co (Lancaster *et al.*, 2011). Les clusters métalliques de la nitrogénase sont synthétisés par les protéines NifUS(Z). Le FeMo-Co nécessite une étape de maturation permise par NifB, puis par le complexe NifEN. L'apport de l'homocitrate et du molybdène est respectivement permise par NifV et NifQ, puis finalement transporté via NafY jusqu'au complexe nitrogénase NifHDK devenant alors fonctionnel (Rubio et Ludden, 2008).

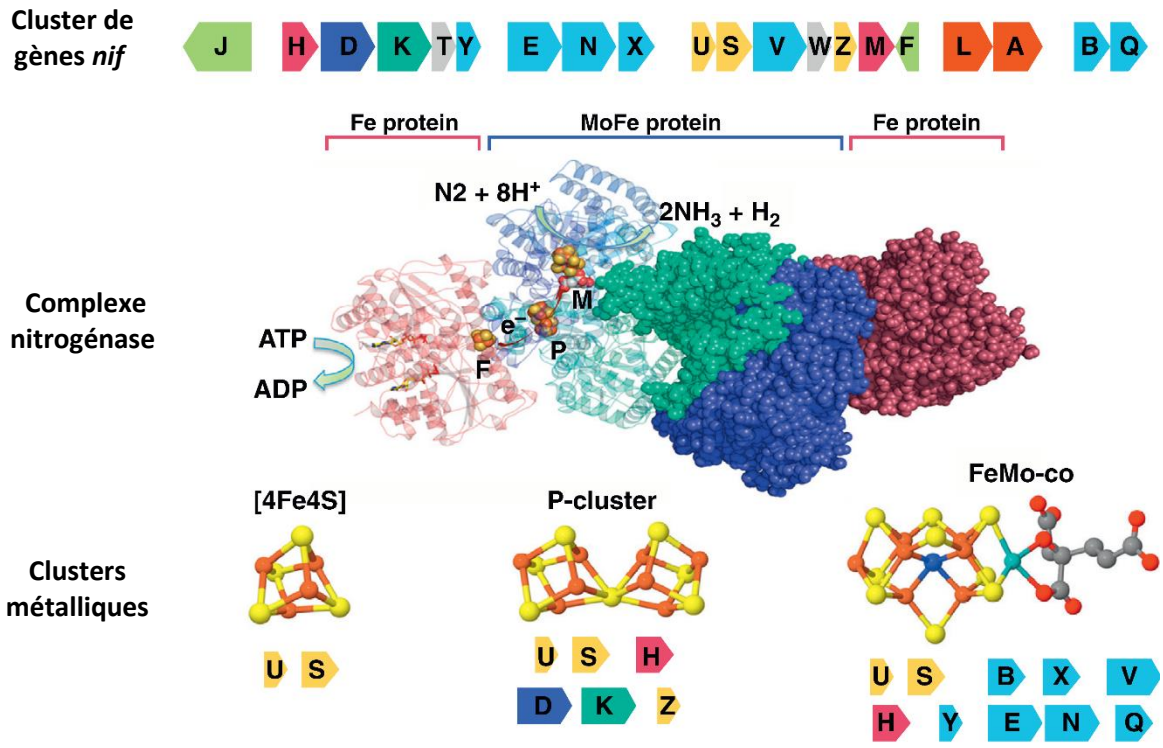


Figure 6 : Organisation génomique et fonctionnelle du complexe nitrogénase. Structure du complexe nitrogénase et gènes nécessaires à sa synthèse. Le cluster *nif* de *Klebsiella oxytoca* est représenté avec un code couleur. Les gènes *nifHDK* sont représentés en couleurs associées à leur structure cristalline. Les autres couleurs représentent les fonctions des gènes, comme ceux en cyan et en jaune, respectivement impliqués dans la synthèse du FeMo-Co et des clusters Fe-S. Les clusters métalliques sont également représentés avec l'ensemble des gènes impliqués dans leur formation (d'après Oldroyd et Dixon, 2014).

La nitrogénase précédemment décrite, dite Mo-Fe, est la plus étudiée. D'autres types de nitrogénases ont été observées dans la nature et sont de type Vanadium-Fer (V-Fe) et type Fer seulement (Fe-Fe), respectivement codées par les gènes *vnf* et *anf*. Ces nitrogénases alternatives sembleraient cependant n'être présentes que chez les organismes possédant également la nitrogénase Mo-Fe. Bien qu'elles soient moins efficaces pour fixer l'azote que la nitrogénase Mo-Fe, il semblerait qu'elles dérivent toutes deux de cette première nitrogénase (Boyd et Peters, 2013). Il existe encore des nitrogénases actuellement non caractérisées parmi les

Archées. Une origine commune entre les protéines responsables de la synthèse des chlorophylles et les nitrogénases est proposée. Une protéine ancestrale NflD pourrait à la fois avoir comme ligand une proto-chlorophylle et un cofacteur NifB. Elle aurait permis respectivement la réduction en chlorophylle ou de l'azote en faible quantité. Des évènements de duplication et de divergence auraient alors finalement donné les protéines BchNB et NifDK actuelles (Boyd et Peters, 2013). La fixation biologique de l'azote ne semble pas avoir été une des caractéristiques de LUCA (Last Universal Common Ancestor, le dernier ancêtre commun de tous les organismes vivants actuels), mais avoir été acquise plusieurs centaines de millions d'années après l'apparition de la vie sur Terre. Aux prémices de la vie, le NH_4^+ fixé de manière abiotique en milieu oxydant devait être suffisant (Boyd et Peters, 2013).

1.3. La diversité des organismes diazotrophes

Les seuls organismes capables de fixer l'azote connus à ce jour sont des Procaryotes (Figure 7). Dans le phyllum des Archées seule les *Euryarchaeota*, sont capables de fixer l'azote. Chez les Eubactéries, les fixateurs sont très répandus, et ce dans quasiment tous les groupes. Au sein d'un même groupe, la capacité de fixation d'azote n'est pas pour autant ubiquitaire. En effet, les phylogénies basées sur l'ARNr 16S regroupent les organismes diazotrophes avec des espèces non fixatrices (Raymond *et al.*, 2004). Cela laisse supposer que la perte secondaire des gènes nécessaires à la fixation d'azote surviendrait de manière récurrente au cours de l'évolution en fonction de la disponibilité en azote de la niche écologique, et/ou que ces gènes seraient acquis par transfert horizontal entre espèces phylogénétiquement éloignées (Boyd et Peters, 2013). Une hypothèse actuelle propose une apparition du complexe nitrogénase chez un ancêtre des Archées méthanogènes il y a -1,5 à -2,2 Ga, bien après l'apparition de la photosynthèse oxygénique (-3,8 Ga) et l'évènement de la Grande Oxydation (-2,5 Ga). Les gènes impliqués dans la fixation d'azote auraient ensuite été transférés horizontalement à un ancêtre des Firmicutes. L'apparition du complexe nitrogénase aurait alors eu lieu dans un environnement anaérobie (Raymond *et al.*, 2004 ; Boyd *et al.*, 2011). Les bactéries du genre *Clostridium* sont les représentants actuels de Firmicutes anaérobies fixatrices d'azote.

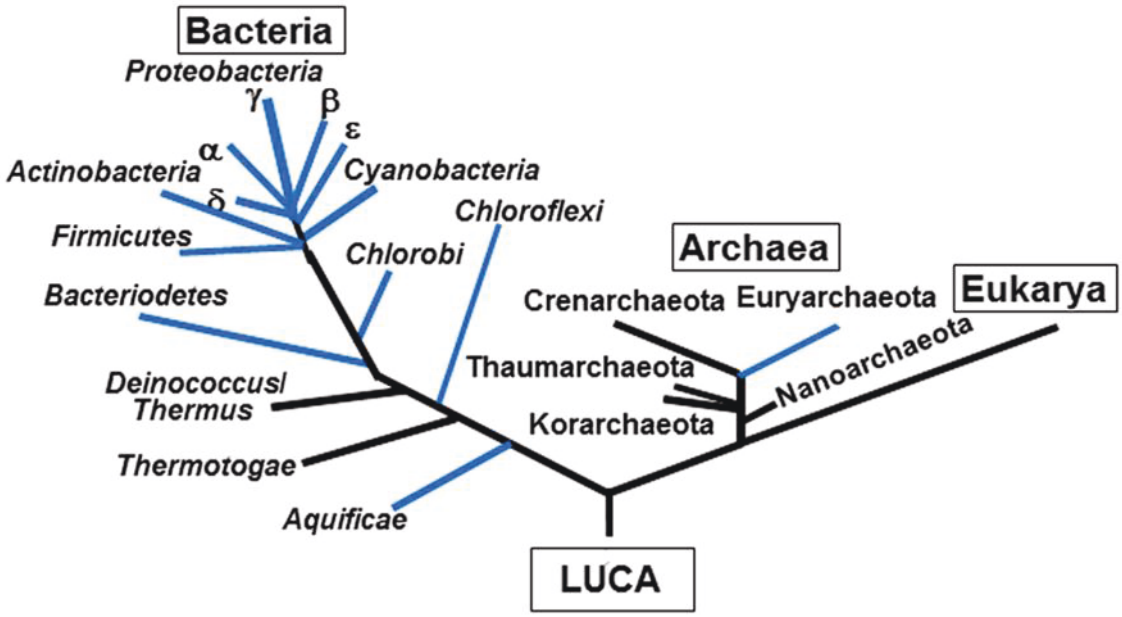


Figure 7: Représentation schématique des trois domaines du vivant, mettant l'accent sur les lignées Procaryotes capables de fixation d'azote. Les lignées diazotrophes sont représentées par des branches bleues, d'après un crible génomique basé sur des séquences d'ADN pour les gènes *nifHDKENB* (d'après Boyd et Peters, 2013).

La capacité de fixation d'azote s'est ensuite étendue aux bactéries anaérobies facultatives comme celles du genre *Klebsiella* ou *Pseudomonas*, ainsi qu'aux bactéries aérobies comme *Acetobacter* et *Azotobacter*. Cependant, l'activité de la nitrogénase est particulièrement sensible à la présence d'O_2 (Fay, 1992). En effet, son activité catalytique est par exemple réduite de moitié en présence de 0,37µM d'O_2 chez *Klebsiella pneumoniae* (Goldberg *et al.*, 1987), et d'une pO_2 de 0,014 atmosphères chez *Azotobacter vinelandii* (Wong et Burris, 1972). Le passage au mode de vie aérobie pour les organismes diazotrophes a nécessité la mise en place de diverses stratégies de réduction de la concentration cellulaire en O_2, évitant ainsi l'inactivation de la nitrogénase. Il existe tout d'abord des adaptations de ségrégation spatiale visibles chez les cyanobactéries filamenteuses du genre *Nostoc* ou *Anabaena* dont les cellules sont interconnectées via des microplasmodesmes. En condition carencées en azote, certaines cellules deviennent des hétérocystes fixateurs d'azote (Figure 8A-D). Cette différentiation morphologique comprend notamment un épaississement de la membrane par accumulation de glycolipides, de lipopolysaccharides, ainsi qu'une limitation du nombre de microplasmodesmes (Mullineaux *et al.*, 2008). De surcroit, la synthèse du photosystème II, permettant la photolyse de l'eau et la libération de O_2, est stoppée et remplacée par celle d'oxydases terminales (Thomas *et al.*, 1970 ; Valladares *et al.*, 2003). Ces modifications ont pour effet de créer un

environnement microoxique, limitant la concentration cellulaire en O_2 (Mullineaux *et al.*, 2008).

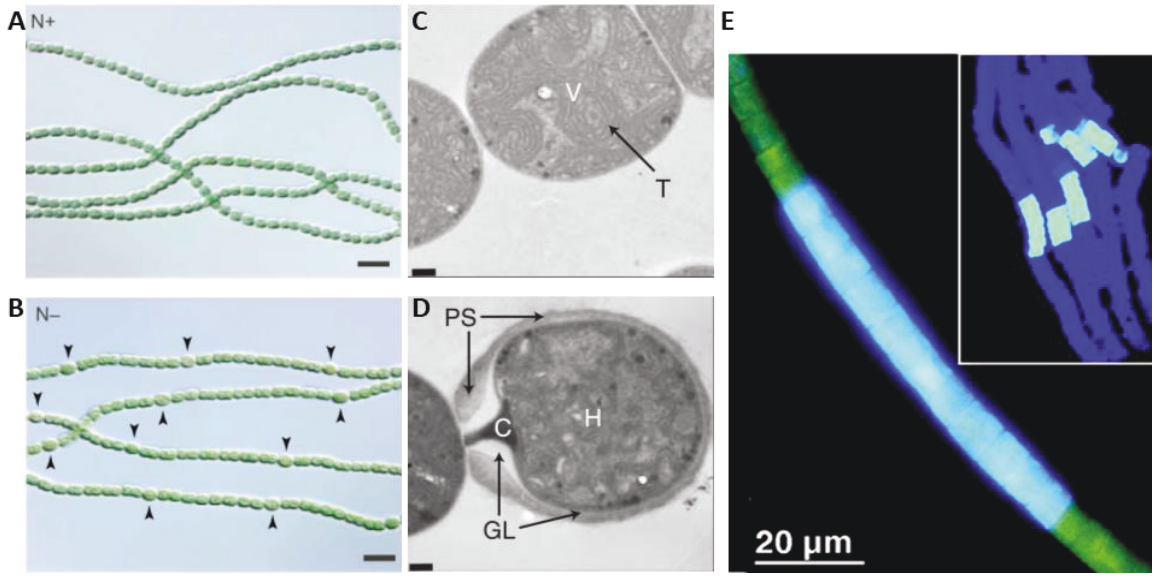


Figure 8: Les cyanobactéries fixatrices d'azote. (A, B) Culture de cyanobactéries *Anabaena* PCC 7120 avec et sans nitrate, respectivement. On observe uniquement des cellules végétatives en présence de nitrate (A) et la formation d'hétérocystes en absence de nitrate signalés par les pointes de flèches (B). Barre d'échelle, 10 µm (d'après Golden et Yoon, 2003). (C, D) Images de microscopie électronique en transmission de cellules végétatives (C) et d'hétérocystes *d'Anabaena* (D); T, membrane de thylakoïdes ; PS, couche de polysaccharide ; GL, couche de glycolipides ; barre d'échelle, 0.2 µm (d'après Kumar *et al.*, 2010). (E) Filaments de *Trichodesmium* IMS101 dont la nitrogénase détectée par immunofluorescence (bleue). L'insert représente une population de *Trichodesmium* marqués de la même manière (d'après Berman-Franck *et al.*, 2001).

Une séparation temporelle est souvent observée parmi les cyanobactéries ne présentant pas de différenciation en hétérocystes. La photosynthèse oxygénique a lieu le jour et la fixation d'azote la nuit, mais cela réduit le taux de fixation d'azote du fait de la plus faible quantité d'ATP en absence de lumière. Chez la cyanobactérie filamenteuse *Trichodesmium* spp., qui est considérée comme l'espèce dominante en termes de fixation d'azote océanique, la fixation d'azote et de la photosynthèse oxygénique est à la fois spatiale et temporelle. L'activité nitrogénase est cantonnée à un petit nombre de cellules (Figure 8E), et suit également un rythme nycthéméral où la lumière du matin induit la photosynthèse oxygénique qui produit des glucides et réduit rapidement le réservoir de plastoquinones. Cet excès de composés réduits réprime l'expression du photosystème II qui combinée à la haute activité respiratoire consommatrice d'O_2 créent un environnement microoxique, induisant de fait la production du complexe nitrogénase et son fonctionnement. La fixation d'azote perdure jusqu'à épuisement des

composés issus de la photosynthèse matinale et l'oxydation progressive des quinones, le taux d'O_2 produit excédant alors celui d'O_2 consommé (Berman-Franck *et al.*, 2001).

Parmi les diazotrophes connus, les bactéries du genre *Frankia* sont très étudiées (phyllum des Actinobactéries). Chez ces actinomycètes filamenteuses, la fixation d'azote a lieu dans des cellules spécialisées appelées vésicules et situées aux extrémités des hyphes (Figure 9 A-B). Le mécanisme limitant la pO_2 cellulaire est différent de celui décrit chez les cyanobactéries puisqu'il consiste en un grand nombre de couches lipidiques (>90 couches) en condition aérobies (Abeysekera *et al.*, 1990). Ces couches lipidiques superposées sont constituées à 80% d'hopanoïdes, des analogues fonctionnels de stérols produits par les Procaryotes (Berry *et al.*, 1993 ; Kleeman *et al.*, 1994 ; Figure 9C). La protection de la nitrogénase contre le dioxygène est un processus dynamique. En effet, il a été observé que plus la pO_2 est élevée, et plus les gènes impliqués dans la synthèse d'hopanoïdes sont induits (Ghodhbane-Gtari *et al.*, 2014).

Il existe également de nombreuses bactéries diazotrophes appartenant aux phyllum des Protéobacteries, et plus particulièrement les protéobactéries du sol portant le nom de rhizobia qui seront plus abondamment décrites dans les chapitres suivants.

Figure 9: Les actinomycètes du genre *Frankia*. (A) Hyphes de *Frankia* avec vésicules fixatrices d'azote observées au microscope à contraste de phase. (B) Vésicule de *Frankia*. Barre d'échelle : 500nm (C) Détail de l'enveloppe à couches multiples d'hopanoïdes d'une vésicule de *Frankia* observée par microscopie électronique en transmission. Barre d'échelle : 50nm (d'après Belin *et al.*, 2018).

1.4. Les symbioses fixatrices d'azote

1.4.1. Interactions symbiotiques et origines des Eucaryotes

Très tôt après l'apparition de la vie, les processus d'associations mutualistes ont eu un impact immense sur la structuration de la biosphère. L'origine même des organismes Eucaryotes, du fait de leur nature hybride, bien qu'encore discutée, peut être proposée comme provenant d'une première endosymbiose entre des ancêtres communs d'Archées et de Eubactéries, respectivement hôte et symbionte (Williams *et al.*, 2013 ; McInerney *et al.*, 2014). La mitochondrie, organite ubiquitaire chez les Eucaryotes, possède un ancêtre commun avec les alphaprotéobactéries (Martijn *et al.*, 2018). L'origine de l'appareil mitotique actuel a également été relié à l'incorporation d'un parasite proche des *Spirochaetae* (Sagan, 1967). La capacité des organismes Eucaryotes à fixer le carbone atmosphérique est obligatoirement passée par un/des processus d'endosymbiose(s) ayant mené à la formation des chloroplastes transmis de façon verticale à la descendance.

L'hypothèse actuelle est que l'origine des chloroplastes remonterait à un évènement unique d'endosymbiose d'une cyanobactérie autotrophe au carbone par un eucaryote hétérotrophe (Zimorski *et al.*, 2014). Cette première lignée d'Eucaryotes photosynthétiques, les Archaeplastida, a ensuite divergé en trois lignées distinctes : Les Glaucophytes, les Rhodophytes et les Chloroplastida. Ces deux dernières lignées se sont d'ailleurs diversifiées et complexifiées au travers d'endosymbioses secondaires voire tertiaires (Gould *et al.*, 2008 ; Zimorski *et al.*, 2014). La fixation biologique de l'azote a elle aussi mené à des symbioses entre des Procaryotes diazotrophes et des hôtes Eucaryotes, avec divers degrés d'intimité observés entre les partenaires. L'hôte échange l'azote assimilable contre une niche écologique favorable au développement des symbiontes et/ou des composés carbonés.

1.4.2. Exemples de symbioses fixatrices d'azote avec des métazoaires

Avant de conquérir les continents, des associations symbiotiques se sont très probablement d'abord formées en milieu marin. Par exemple, certains microorganismes fixateurs d'azote sont des symbiontes connus de divers invertébrés actuels comme les mollusques bivalves, les éponges, les coraux, les oursins et mêmes des tuniciers (Fiore *et al.*, 2010 ; Petersen *et al.*, 2016). De nombreux protistes tels que les dinoflagellés, les diatomées, et les radiolaires en

possèdent également (Villareal, 1992 ; Fiore *et al.*, 2010). Les symbiontes fixateurs d'azote et cellulolytiques présents dans la glande de Deshayes des *Bankia setacea* (Lamellibranche) sont transmis verticalement (Sipe *et al.*, 2000). Chez certaines diatomées, une transmission verticale de cyanobactéries symbiotiques, les *Richelia intracellularis*, a également été observée (Janson *et al.*, 1999).

1.4.3. Symbioses fixatrices d'azote entre cyanobactéries et Embryophytes

A partir du Dévonien (-420 Ma) jusqu'à nos jours, les Embryophytes ont conquis le milieu terrestre devenant ainsi les organismes contribuant majoritairement à l'accumulation de biomasse sur les continents. Leur importance écologique est immense car ils sont à la base des réseaux trophiques terrestres. Ces organismes se sont diversifiés avec leurs symbiontes, permettant alors de coloniser de nouvelles niches écologiques. Les cyanobactéries du genre *Nostoc* et/ou *Anabaena* peuvent coloniser de façon intercellulaire des espèces d'*Anthoceros* Bryophytes au niveau de cryptes situées sur le thalle du gamétophyte (Figure 10A). Les cavités présentes dans les jeunes feuilles de Ptéridophytes peuvent également être colonisées par ces cyanobactéries diazotrophes (Figure 10B). L'association entre les fougères aquatiques du genre *Azolla* et les cyanobactéries diazotrophes des genres *Anabaena* et *Nostoc* en est un exemple agronomiquement important car utilisé comme engrais vert dans les rizières (Santi *et al.*, 2013). C'est également une piste explorée pour la nutrition des animaux d'élevage, notamment dans des zones où la productivité des cultures agricoles classiques est faible (Rawat *et al.*, 2015). Les Gymnospermes de l'ordre des *Cycadales* et les Angiospermes du genre *Gunnera* sont également colonisées par des cyanobactéries endophytes (Figure 10C-D). Ces dernières sont présentes dans des racines adventives dites coralloïdes à la base du stipe des *Cycadales*, formant un cercle bleu-vert dans la zone corticale, et dans des glandes présentes sur la tige à la base des feuilles dans le cas de *Gunnera*. Les plantes hôtes induisent via un signal chimique la transformation des cyanobactéries en hormogonies, qui en sont la forme motile et capable de pénétrer dans les cavités ou tissus de la plante. La production de ces structures d'accueil est contrôlée et dépendante de la biomasse de la plante, exerçant ainsi un contrôle sur les populations de cyanobactéries (Rai *et al.*, 2000). Une fois les cyanobactéries internalisées dans les cavités des frondes, les plantes induisent un nombre d'hétérocystes supérieur à celui observé en vie libre, en leur fournissant du carbone fixé par la photosynthèse eucaryote, augmentant alors le ratio C/N de leur environnement (Rai *et al.*, 2000 ; Meeks et Elhai, 2002).

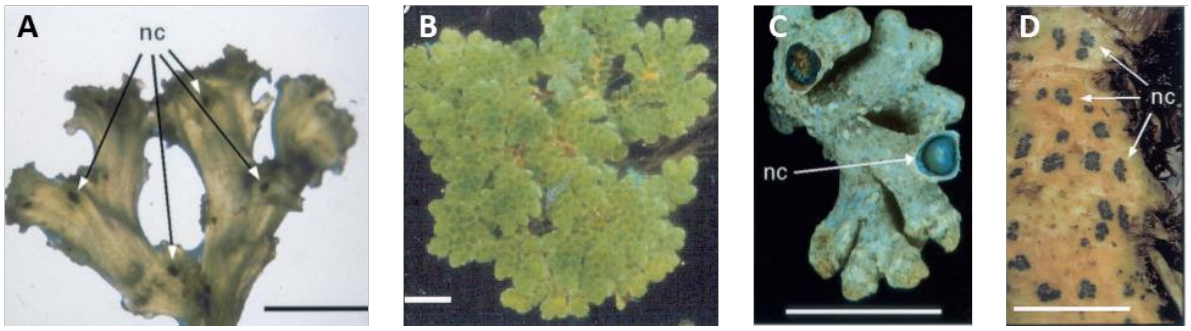


Figure 10: Les symbioses fixatrices d'azote entre Embryophytes et cyanobactéries. Les structures colonisées par des cyanobactéries (*Nostoc* ou *Anabaena*) fixatrices d'azote sont représentées par des flèches. (A) Thalle *d'Anthoceros* spp. (B) Le sporophyte d'*Azolla* spp. (C) Coupe de racine coralloïde de l'espèce *Cycas taiwaniana* de la famille des Cycadales. (D) Cyanobactéries dans les glandes caulinaires de *Gunnera chilensis*. nc : colonies de *Nostoc*. Barres d'échelle : A : 1cm, B : 0,25cm, C : 0,5cm, D : 1cm (d'après Meeks et Elhai, 2002).

1.4.4. Symbioses fixatrices d'azote entre Angiospermes et bactéries rhizosphériques ou endophytes

Les végétaux terrestres sont également capables de modifier les communautés microbiennes du sol les environnant. Ce contrôle s'effectue notamment par la libération d'exsudats racinaires potentiellement toxiques et/ou métabolisables, provoquant ainsi une sélection de la communauté microbienne. Les espèces résistantes ou adaptées à la niche écologique ainsi produite sont favorisées dans la rhizosphère, définie comme la zone du sol influencée par les racines (Baetz et Martinoia, 2014). Parmi les communautés bactériennes rhizosphériques, les *Plant Growth-Promoting Rhizobacteria* (PGPR) ou rhizobactéries favorisant la croissance des plantes intéresse depuis longtemps la communauté scientifique. Ces dernières amélioreraient la croissance végétale via divers mécanismes d'intérêt nutritionnels et/ou immunitaires (Bhattacharyya et Jha, 2012). L'un d'entre eux est la fixation biologique de l'azote, réalisée par la plupart des PGPR. Il existe notamment des associations endophytes entre des protéobactéries diazotrophes et des plantes de la famille de Poacées telles que le riz, le maïs, le blé et la canne à sucre. Cette dernière peut être fertilisée par l'activité nitrogénase de *Gluconacetobacter diazotrophicus* à raison de 200kg d'azote par hectare (Eskin *et al.*, 2014), contribuant jusqu'à 59% de sa biomasse azotée (Taulé *et al.*, 2012). Mais malgré la présence confirmée des gènes *nifHDK* dans les génomes de nombreuses bactéries endophytes et PGPR, la fixation d'azote n'est pas toujours observée ou l'est seulement dans des conditions particulières. C'est le cas de la bactérie endophyte des racines de riz *Serratia marcescens*, qui semble avoir une activité

PGPR avérée mais dont la fixation d'azote n'est détectée qu'en présence de composés carbonés ajoutés au niveau des racines (Gyaneshwar *et al.*, 2001).

1.4.5. Les interactions endosymbiotiques entre Protéobactéries et Angiospermes

Parmi ces PGPR, certaines sont capables d'associations avec un degré d'intimité maximal avec certaines espèces de plantes hôtes, à l'intérieur même des cellules de ces dernières. Ce type d'association est responsable de la formation d'un nouvel organe, la nodosité, au niveau des racines (et plus rarement des tiges) de la plante hôte. Pour des raisons de simplicité, nous utiliserons les anglicismes/néologismes « nodulation », désignant le processus de mise en place de nodosités ; et « noduler », l'aptitude d'une souche bactérienne à initier des nodosités, respectivement. Les cellules des nodosités sont infectées par des bactéries fixatrices d'azote, alors appelés bactéroïdes une fois internalisés dans les cellules végétales, via des mécanismes qui sont présentés dans le chapitre suivant. Ce type de symbiose ne concerne que certains groupes de plantes Angiospermes en association avec l'actinomycète *Frankia* ou avec des protéobactéries à coloration de Gram négative, appelées rhizobia.

Les *Frankia* réalisent des symbioses avec des plantes actinorhiziennes, dont les nodosités racinaires fixatrices d'azote sont appelées actinorhizes. Ces plantes se répartissent en 25 genres et 8 familles d'Angiospermes appartenant à trois ordres principaux : les Fagales (ex : *Casuarina*, *Alnus*), les Cucurbitales (ex : *Datisca, Coriaria*) et les Rosales (ex : *Elaeagnus*, *Discaria*) (Svistoonoff *et al.*, 2014). Les rhizobia quant à elles réalisent des symbioses uniquement avec des plantes de la famille des *Fabaceae* également appelées légumineuses. Une exception a été observée : de nombreuses espèces de rhizobia peuvent induire des nodosités fixatrices d'azote chez *Parasponia* spp., un genre appartenant à l'ordre des Rosales et à la famille des *Cannabaceae* (Op den Camp *et al.*, 2012). Cette symbiose était considérée comme récemment apparue et utilisée comme un modèle pour l'étude des phases précoces de mise en place des nodosités. Cependant, des études de génomique comparative récentes laissent penser que cette symbiose pourrait être ancienne. De plus, l'ancêtre commun des *Parasponia* spp. et des autres *Cannabaceae* non symbiotiques telles que *Trema tomentosa* aurait été capable de nodulation. En effet, l'absence de capacités symbiotiques de la majorité des *Cannabaceae* serait dû à des pertes secondaires de gènes essentiels à l'établissement des nodosités. Cela soulève alors la possibilité que la nodulation trouverait son origine chez un ancêtre commun de toutes

les plantes capables de mettre en place des nodosités il y a environ 100 Ma (van Velzen *et al.*, 2018).

Il est important de noter que dans le cadre de cette symbiose nutritionnelle, le même problème d'inactivation de la nitrogénase par l'oxygène se pose que lors de la fixation libre. Les vésicules de *Frankia* dans les actinorhizes d'*Alnus* sont, comme en condition de fixation d'azote en vie libre, protégées par de très nombreuses couches d'hopanoïdes. Cependant, chez les légumineuses, les *Parasponia* en symbiose avec les rhizobia ainsi que chez les *Casuarina* associée avec les *Frankia*, une autre stratégie est mise en place (Appleby *et al.*, 1983 ; Gherbi *et al.*, 1997). Elle implique une famille de protéines présente de façon ubiquitaire chez les Eucaryotes que l'on appelle les hémoglobines (Dordas *et al.*, 2003). Connues pour leur fonction dans le transport de l'O_2 chez les vertébrés, elles existent également chez les organismes végétaux en conditions non-symbiotiques et symbiotiques. Les hémoglobines symbiotiques telles que la leghémoglobine des légumineuses sont les protéines végétales les plus abondantes des nodosités. Elles ont pour rôle de fixer le dioxygène, puis de l'apporter progressivement aux symbiontes en le séquestrant, pour améliorer l'activité de fixation de l'azote (Udvardi et Poole, 2013). Le rôle écologique joué par ce type d'interaction symbiotique est majeur sur les continents. Elle permet d'enrichir le sol en azote et d'améliorer la résistance des végétaux aux stress nutritionnels. De fait, les plantes réalisant ces symbioses sont des espèces pionnières dans la colonisation d'un nouvel environnement. L'ensemble des associations symbiotiques existantes entre bactéries réalisant la fixation biologique de l'azote et Embryophytes est résumé ci-dessous (Mus *et al.*, 2016 ; Figure 11).

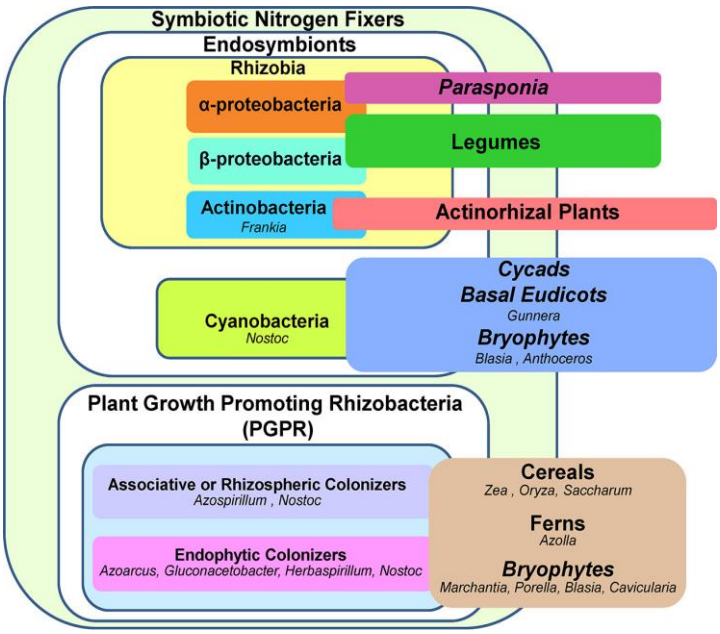


Figure 11: Représentation des associations symbiotiques possibles entre bactéries réalisant la fixation biologique de l'azote et Embryophytes (d'après Mus *et al.*, 2016).

2. Les symbioses rhizobium-légumineuses

2.1. Généralités sur les légumineuses

2.1.1. Diversité et écologie des légumineuses

On estime que la famille des *Fabaceae* (ou légumineuses) regroupe environ 750 genres comprenant presque 20 000 espèces. En 2017, une nouvelle classification des sous-familles de *Fabaceae* basée sur la phylogénie de 3842 séquences du gène chloroplastique *matK* a été publiée (Azani *et al.*, 2017). Cette nouvelle classification compte 4 nouvelles sous-familles de *Fabaceae* non-nodulées qui sont les *Cercidoideae* (arbres, arbustes et lianes), *Detarioideae* (arbres et arbustes), *Duparquetioideae* (lianes) et les *Dialioideae* (arbres et arbustes) (Figure 12). Les membres des sous-familles des *Caesalpinioideae* et des *Papilionoideae* peuvent être capables de nodulation, et la sous famille des *Papilionoideae* est la plus diversifiée au sein des *Fabaceae*. Elle représente environ 14000 espèces réparties de façon cosmopolite sur les continents, témoignant ainsi de leur capacité à s'adapter à un grand nombre d'environnements différents (Azani *et al.*, 2017).

La diversification des légumineuses a démarré il y a 60 Ma alors qu'il n'existait que deux grands continents, la Laurasia et le Gondwana séparés par la Téthys équatoriale. Les mouvements des plaques tectoniques ont mené au déplacement et à la séparation de masses continentales. Ceci a provoqué des modifications des conditions abiotiques, diversifiant les environnements conquis par les légumineuses (Sprent *et al.*, 2013). Cette diversité de biomes conquis par les légumineuses est due en grande partie dans leur capacité à être nodulées, c'est-à-dire à effectuer une symbiose fixatrice d'azote avec les rhizobia. Cela leur permet d'être des espèces pionnières du fait de la possibilité de coloniser des milieux pauvres en azote. Les biomes tropicaux humides sont colonisés par les légumineuses, avec une dominance des *Caesalpinioideae* par rapport aux *Papilionoideae*. Les *Caesalpinioideae*, bien que réparties également sur tous les continents, sont plus diversifiées en Amérique et présentent un point chaud de diversité au Brésil. C'est aussi dans cette région qu'il existe le plus de *Caesalpinioideae* capables de nodulation, là où celles des autres régions du monde en ont perdu la capacité (Sprent *et al.*, 2017). De nombreuses espèces d'*Acacia* (Mimosoïdes, appartenant aux *Caesalpinioideae*) sont présentes dans les régions semi arides Africaines et tolérantes à la sécheresse, capables ou non de nodulation (Sprent *et al.*, 2010). La flore Australienne native est particulièrement riche en légumineuses, du fait de la diversité de ses biomes existants, ainsi que de l'existense de deux évènements indépendants de colonisation au cours des temps

géologiques (Sprent *et al.*, 2017). Le genre *Glycine* s'y est particulièrement diversifié, même si l'espèce la plus connue, *Glycine max* (le soja) est originaire de Chine (Sprent *et al.*, 2017).

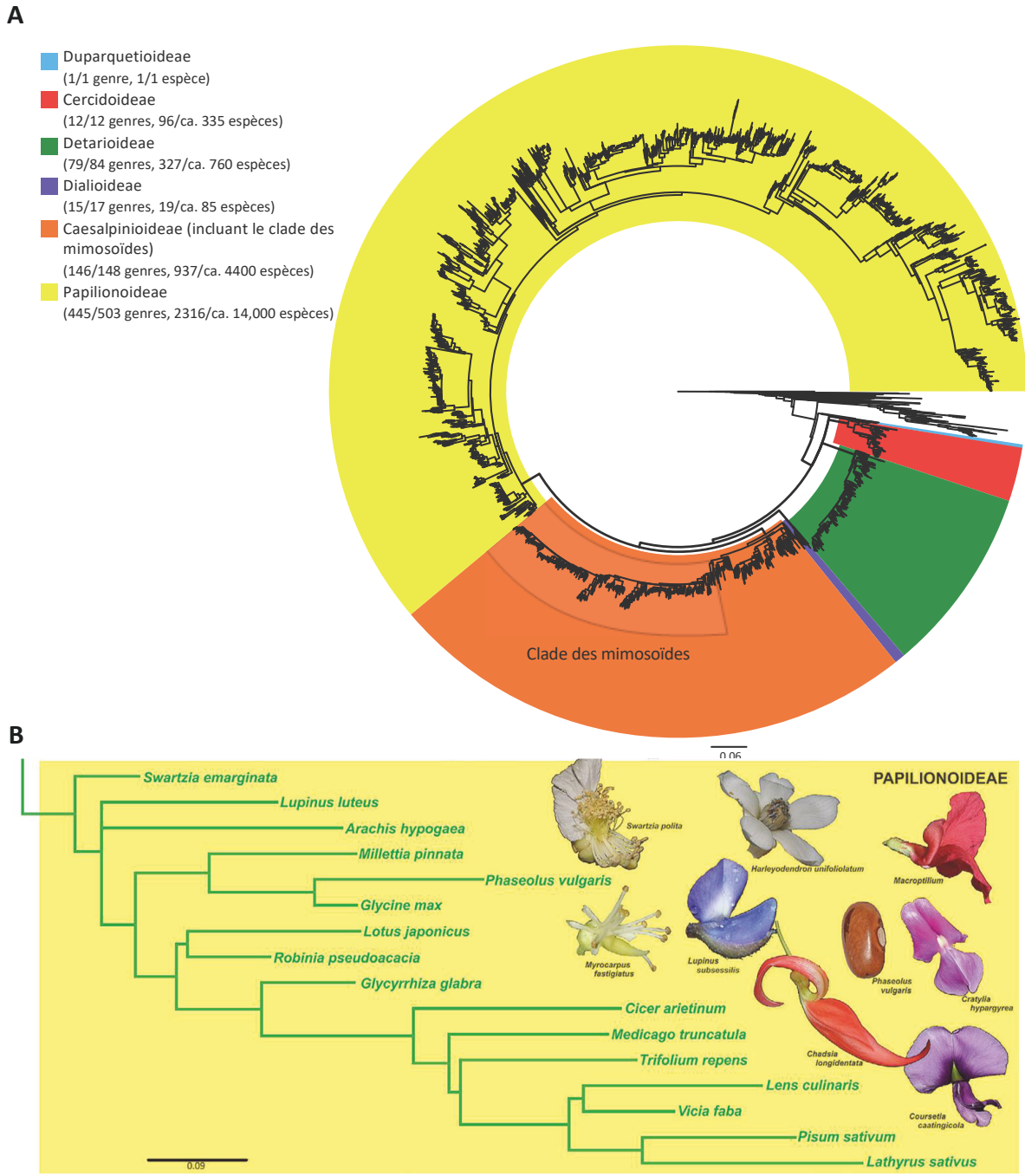


Figure 12: Nouvelle classification des *Fabaceae* basée sur la phylogénie des séquences *matK*. (A) Arbre phylogénétique consensus basé sur 3842 séquences de gènes *matK*. La longueur des branches est proportionnelle aux nombre de substitutions chez les différents *matK*. (B) Détails de l'arbre pour des espèces appartenant à la sous famille des Papilionoideae (d'après Azani *et al.*, 2017).

Parmi la sous famille des *Papilionoideae*, il existe le clade monophylétique des IRLC (pour Inverted Repeat Lacking Clade) défini par une mutation structurale caractéristique : la perte d'une région du génome chloroplastique de 25kb (kilobases) (Wojciechowski *et al.*, 2000). Les IRLC seraient apparues il y a 39 Ma et elles constituent un des plus grands clades des *Fabaceae* avec plus de 4500 espèces, dont 2300 rien que pour le genre *Astragalus* (Sprent *et al.*, 2013). Certains membres de ce clade sont naturellement présents des latitudes les plus élevées du cercle polaire arctique aux régions méditerranéennes (Sprent *et al.*, 2017). Les plantes du genre *Medicago*, *Trifolium*, *Astragalus* ont colonisé les toundras polaires et les régions de haute altitude comme le plateau du Tibet. La symbiose fixatrice d'azote avec les rhizobia peut s'établir dans quasiment tous les biomes colonisés par les légumineuses. Cette interaction spécifique entre les deux partenaires Eucaryote et Procaryote revêt de nombreux intérêts agronomiques.

2.1.2. Importance agronomique des légumineuses

Les *Fabaceae* ou légumineuses font partie intégrante de l'alimentation humaine de manière directe, ou indirecte via l'alimentation des animaux d'élevage. Les légumineuses comme le haricot, le pois, la lentille ou l'arachide sont directement consommées pour leurs graines. Le bétail peut être nourrit avec des légumineuses fourragères (luzerne, trèfle) avec des mélanges de soja et de graminées ou des coproduits de leurs transformation (tourteaux). Les légumineuses cultivées appartiennent surtout à la sous famille des *Papilionoideae*, dont 97% des espèces étudiées sont capables d'être nodulées par des rhizobia fixatrices d'azote (Bruneau *et al.*, 2013). L'azote n'est donc pas pour ces plantes un élément limitant leur croissance et leur capacité de reproduction. Les légumineuses ont des parties aériennes et racinaires riches en protéines qui enrichissent le sol une fois labourées, et certaines espèces produisent des graines consommables riches en protéines (protéagineux). Des espèces comme *Trifolium subterraneum* ont été introduites en Australie et utilisées comme plantes fourragères pour le bétail dès les années 1830. Certains auteurs estiment que l'apport écosystémique de la luzerne cultivée *Medicago sativa*, originaire de Chine et exportée partout dans le monde, contribue à plus de mille milliards de dollars de revenus dans l'économie (Sprent *et al.*, 2017).

Dans le monde et notamment en France, la culture des légumineuses était très employée du fait de leur capacité à se développer sans intrants chimiques azotés sur des sols pauvres et de

leur utilisation dans la rotation des cultures pour réenrichir le sol en azote. A partir des années 1960 s'est déroulée la « révolution fourragère » dont l'objectif était de répondre à une forte demande de l'après-guerre en produits agricoles et plus particulièrement en produits carnés. Dans ce contexte, les légumineuses ont été remplacées progressivement par des cultures de graminées à haut rendement et abondamment fertilisées par des engrais azotés aux prix, à l'époque, attractifs (Voisin *et al.*, 2013). En effet, les rendements mondiaux en légumineuses à graines n'ont progressé que de 550 kg/ha en 1961 à 1 000 kg/ha en 2013 quand celle des rendements céréaliers sont passés de 1500 kg/ha à 4 000 kg/ha au cours de la même période (Joshi et Rao, 2017 ; Figure 13). Les effets précédemment décrits des engrais azotés ainsi que l'augmentation de leur prix ont relancé l'attractivité des cultures de légumineuses fourragères et à graines. En France et en Europe, la culture du pois est majoritaire parmi les légumineuses consommées. Elle peut atteindre des rendements en France oscillant entre 3 500 et 5 000 kg/ha, loin derrière les rendements de blé tendre dépassant les 7 000 kg/ha (source : UNIP).

Les graines de légumineuses sont une source importante et inévitable de protéines pour les personnes ne consommant pas ou peu de produits d'origine animale. Par conséquent, l'Inde, qui compte environ 30% de végétariens dans sa population, en est de très loin le plus grand producteur et importateur (Joshi et Rao, 2017). La surface des terres arables où des légumineuses à graines sont cultivées dépasse les 34% ($2{,}7.10^7$ ha), pour une production représentant 25% de la production végétale mondiale. A l'échelle mondiale, les graines de légumineuses les plus produites sont celles de soja (*Glycine max*) avec des débouchés pour l'alimentation humaine et celle du bétail, mais aussi comme source de biocarburants (Masuda et Goldschmidt, 2009). La production mondiale, dominée par les Etats-Unis, le Brésil et l'Argentine, est de plus de 347 millions de tonnes pour l'année 2017-2018 (source : sopa.org).

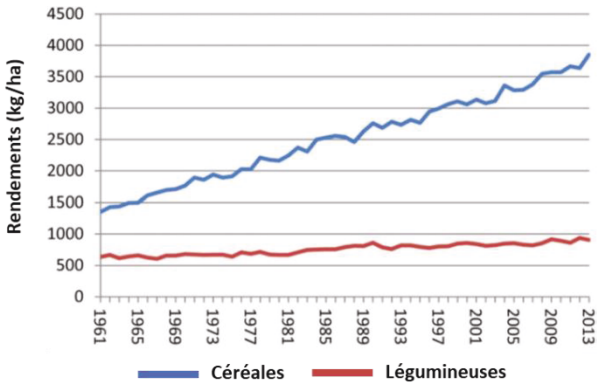


Figure 13: Evolution des rendements mondiaux en céréales et en légumineuses à graines de 1961 à 2013 (d'après Joshi et Rao, 2017).

2.2. La rhizosphère et ses habitants

2.2.1. La rhizosphère, un lieu d'interactions entre plantes et microorganismes

Le sol est extrêmement riche en microorganismes divers. Il existerait jusqu'à 10 000 espèces bactériennes et près de 10^9 bactéries cultivables par gramme de sol (Poole *et al.*, 2018). La composition microbienne du sol dépend de nombreuses conditions abiotiques (pH, température) et biotiques. En particulier, les espèces végétales exercent une influence considérable sur la structure du microbiote du sol. Cet effet est croissant depuis la zone du sol influencée par les racines appelée rhizosphère jusqu'à l'interface sol/racine appelée rhizoplan (Poole *et al.*, 2018). La biodiversité microbienne rhizosphérique est en générale plus faible que celle du sol mais la quantité de bactéries y est souvent plus forte, témoignant de la sélection effectuée par les plantes qui favorisent la prolifération spécifique de quelques taxons (Poole *et al.*, 2018).

Rapidement après l'imbibition de leurs graines, certaines espèces de plantes libèrent des molécules dérivant du métabolisme secondaire, appelées flavonoïdes. Les rôles des flavonoïdes excrétés sont multiples, mais il en existe qui ont des propriétés chimioattractantes sur certaines espèces bactériennes, participant ainsi en partie à la sélection du microbiote rhizosphérique. C'est notamment le cas des bactéries de l'espèce *Ensifer meliloti* (anciennement *Sinorhizobium meliloti*) qui sont attirées par la lutéoline (ie. un flavonoïde) produite par *Medicago truncatula* en cours de germination. La lutéoline induit également la formation de biofilms bactériens au niveau rhizosphérique et une plus grande motilité bactérienne (Liu et Murray, 2016 ; Figure 14). Les flavonoïdes ont également pour particularité d'induire le transfert d'éléments génétiques mobiles requis pour la nodulation et la fixation d'azote entre bactéries compatibles décrits ci-dessous (Ling *et al.*, 2016 ; Figure 14). Les flavonoïdes sont impliqués dans le dialogue moléculaire menant à la formation de nodosités, dont les mécanismes sont détaillés plus loin (Figure 14).

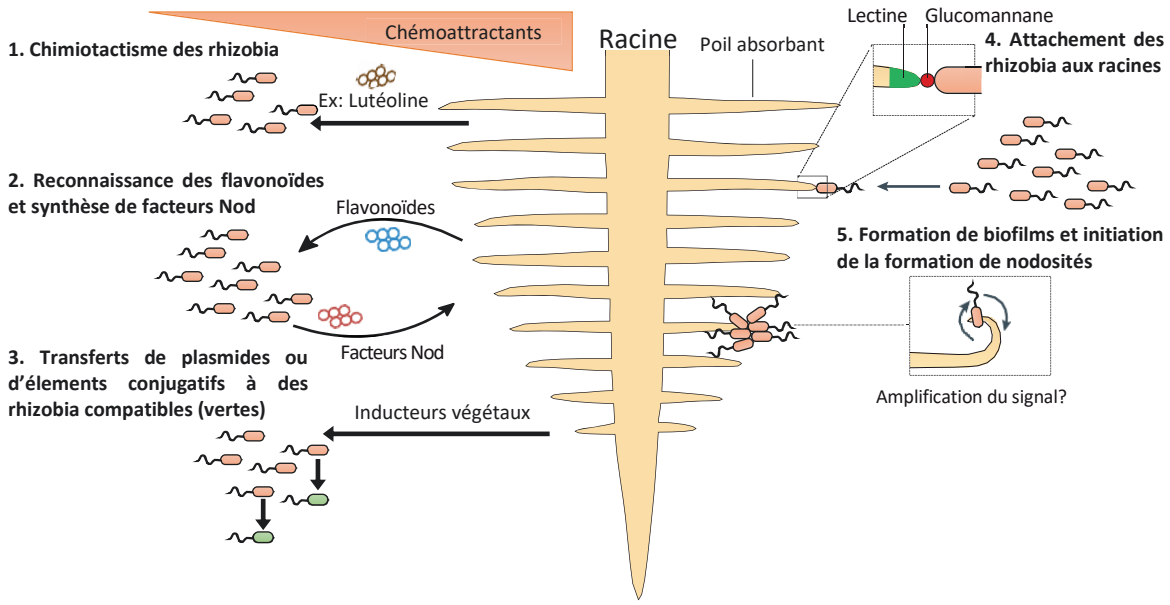


Figure 14: Attachement et colonisation des racines de légumineuses par les rhizobia. Cinq types d'interactions entre les rhizobia et les légumineuses sont présentés. Les bactéries ne sont pas représentées à l'échelle pour des raisons de visibilité (modifié d'après Poole *et al.*, 2018).

2.2.2. La structure génomique des rhizobia et des gènes symbiotiques

Le terme de rhizobia désigne un groupe polyphylétique. En effet il regroupe à la fois des α- et des β-protéobactéries (Masson-Boivin *et al.*, 2009). Leur point commun est leur capacité à établir une symbiose fixatrice d'azote avec des Angiospermes où l'interaction se déroule dans des nodosités racinaires (et parfois caulinaires). Les génomes des rhizobia sont en général d'assez grande taille (entre 5 et 10 Mb) et présentent souvent une architecture complexe formée d'un chromosome et de plasmides additionnels de tailles variables (Poole *et al.*, 2018).

La capacité d'établir une interaction symbiotique avec les légumineuses nécessite en premier ordre de posséder des gènes symbiotiques, dont il existe dans la grande majorité des cas, deux groupes fonctionnels. Le premier groupe de gènes code un complexe nitrogénase fonctionnel tels que les gènes *nif* et *fix* par exemple. Les autres gènes requis sont appelés les gènes *nod*, qui codent des protéines responsables de la biosynthèse de molécules signal qui seront décrites dans le sous-chapitre suivant. L'accumulation des données de séquençages de génomes de rhizobia obtenus à partir des années 2000 a permis de découvrir que ces gènes (*nif*, *fix*, *nod*) sont regroupés dans des zones précises du génome appelés îlots symbiotiques. Ils sont également présents au sein d'éléments mobilisables (ou ICE pour *Integrative and Conjugative*

Elements), pouvant être transférés horizontalement à d'autres microorganismes (Poole *et al.*, 2018, Figure 15).

Les gènes symbiotiques sont soit présents sur des plasmides symbiotiques comme chez *E. meliloti* (Galibert *et al.*, 2001), ou bien intégrés dans le chromosome bactérien sous forme d'îlot(s) symbiotique(s) comme chez *Mesorhizobium loti* et *Bradyrhizobium diazoefficiens* USDA110 (Kaneko *et al.*, 2000 ; Kaneko *et al.*, 2002). Les îlots symbiotiques font partie d'une classe plus large d'éléments génétiques mobiles appelés îlots génomiques, qui sont des régions d'ADN plus ou moins grandes augmentant la valeur sélective d'une bactérie dans une niche écologique donnée. Ces îlots sont souvent bornés par des séquences d'ARNt, séquences d'intégrations les plus communes car très conservées parmi les Procaryotes. Ils contiennent des gènes codant des intégrases requises pour l'intégration de l'îlot au niveau d'une séquence d'ARNt qui se dédouble de façon asymétrique, et des transposases catalysant l'excision de l'îlot. Ce dernier peut ensuite être transféré à d'autres bactéries (Dobrindt *et al.*, 2004). Les îlots symbiotiques peuvent être également séparés dans différentes régions du chromosome mais être tout de même remobilisés en un unique plasmide conjugatif transférable à d'autres Procaryotes (Haskett *et al.*, 2016).

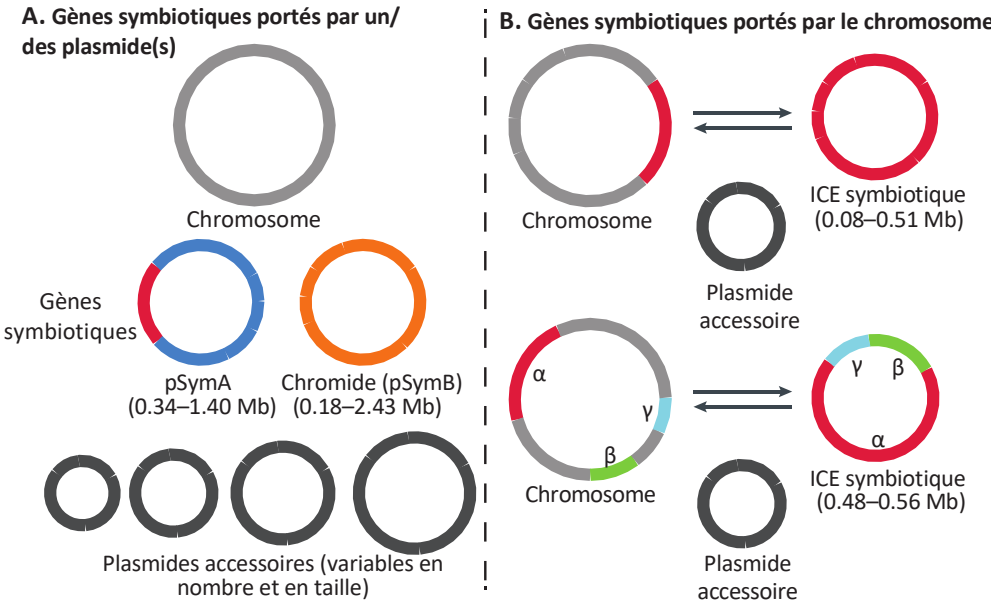


Figure 15: Organisation génomique des rhizobia. Les gènes symbiotiques (*nif*, *fix*, *nod*) peuvent être portés par des plasmides ou intégrés dans le chromosome bactérien (B). (A) L'organisation en multiples réplicons est observée chez les genres *Rhizobium*, *Ensifer*, *Cupriavidus* et *Burkholderia*. Certains plasmides finissent même par devenir essentiels à la survie de la bactérie et sont alors nommés chromides. C'est le cas du plasmide pSymB de *E. meliloti*. (B) Les îlots symbiotiques s'intégrant au niveau des séquences conservées d'ARNt sont composés des gènes symbiotiques ainsi que séquences répétées et de gènes codant des intégrases et transposases. Cette organisation est observée dans les genres *Mesorhizobium*, *Azorhizobium* et *Bradyrhizobium* (d'après Poole *et al.*, 2018).

Chez *E. meliloti*, il existe deux plasmides symbiotiques pSymA et pSymB, qui sont respectivement accessoire et essentiel pour la prolifération de la bactérie en vie libre. C'est la translocation des gènes essentiels tRNA-Arg et *engA* initialement chromosomiques sur le plasmide pSymB qui a mené à sa transformation en chromide, c'est-à-dire plasmide requis pour la survie de la bactérie (DiCenzo *et al.*, 2013 ; Poole *et al.*, 2018). Le reste du génome des rhizobia, appelé core-génome, est très proche de bactéries rhizosphériques, qu'elles soient des PGPR ou des agents pathogènes telles que les genres *Agrobacterium* ou *Ralstonia* (Masson-Boivin *et al.*, 2009, Figure 16). Dès lors, on peut affirmer que l'histoire évolutive des rhizobia est riche de transferts horizontaux de gènes symbiotiques. Un scénario possible d'adaptation de bactéries rhizosphériques devenant fixatrices d'azote symbiotiques après transfert horizontal d'un plasmide symbiotique a été effectuée en laboratoire avec succès via un dispositif d'évolution expérimentale (Remigi *et al.*, 2014). La souche pathogène de plante *Ralstonia solanacearum* GMI1000 a artificiellement intégré le plasmide symbiotique de *Cupriavidus taiwanensis*. La bactérie transformée a subi une pression de sélection en présence de l'hôte naturel de *C. taiwanensis*, *Mimosa pudica*. Cette transition entre pathogène et symbionte de plante est dans ce cas passée par un remodelage génomique transitoire. Ce remodelage est permis par la cassette *imuABC* présente sur le plasmide et codant une ADN polymérase peu fidèle. L'hypermutabilité induite par cette enzyme permet alors l'adaptation rapide à l'hôte végétal (Remigi *et al.*, 2014).

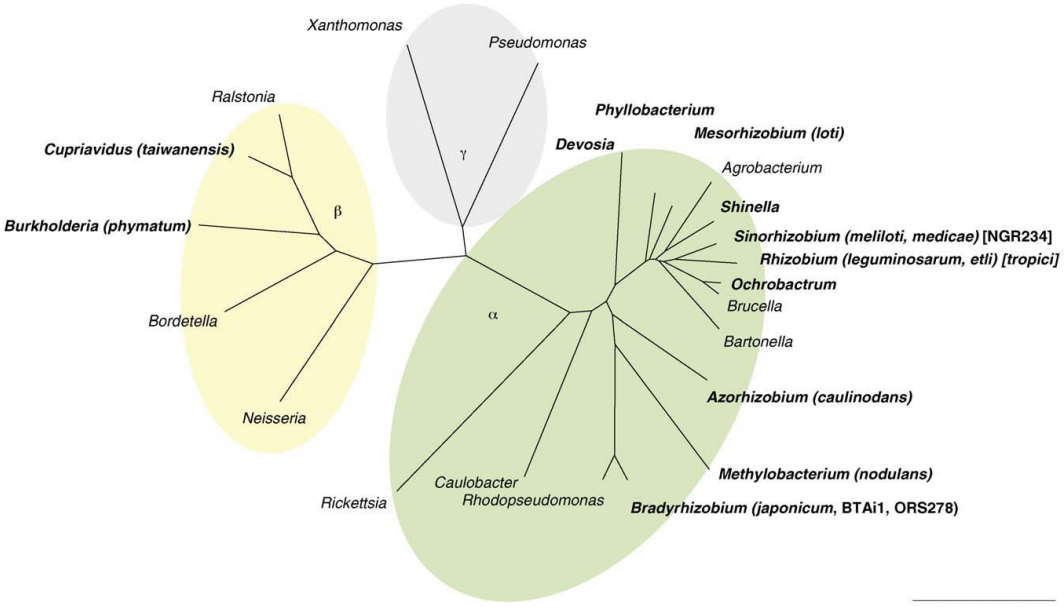


Figure 16: Arbre phylogénétique de séquences d'ADNr 16S provenant d'α-, β- et γ-protéobactéries. Les genres sélectionnés interagissent avec divers degrés d'intimités avec des organismes végétaux. Les genres en caractères gras contiennent des rhizobia, et les noms entre parenthèses indiquent ceux des souches dont le génome a été entièrement séquencé (d'après Masson-Boivin *et al.*, 2009).

2.2.3. La diversité des rhizobia et la variabilité de leur spectre d'hôte

L'importance que revêt le microbiote du sol pour la fixation d'azote a récemment été indirectement démontrée, en observant que la symbiose limite l'établissement des légumineuses en dehors de leur aire de répartition originelle (Simonsen *et al.*, 2017). Les légumineuses non symbiotiques semblent plus adaptées pour coloniser un nouvel environnement comparé aux légumineuses symbiotiques parce qu'elles ne sont pas autant dépendantes vis-à-vis de la communauté microbienne du sol du nouvel environnement (Simonsen *et al.*, 2017). En effet, si les légumineuses symbiotiques ont toujours bénéficié d'azote provenant de leur(s) symbiontes endémique(s), il est possible d'imaginer un relâchement de la sélection des gènes végétaux impliqués dans l'import d'azote de façon non symbiotique. Il se trouve que l'établissement du processus symbiotique résulte d'une coévolution entre hôte et symbiontes qui peuvent devenir plus ou moins spécifiques. De fait, si ces mêmes hôtes changent d'aire de répartition, où les souches de symbiontes avec lesquels ils ont co-évolués sont absents, alors l'établissement de la symbiose sera inefficace voire impossible.

Les rhizobia se répartissent dans un ensemble de 16 genres parmi les α- et β-protéobactéries. Les genres *Rhizobium*, *Pararhizobium*, *Neorhizobium*, *Agrobacterium*, *Ochrobactrum*, *Ensifer*/*Sinorhizobium*, *Phyllobacterium*, *Mesorhizobium*, *Aminobacter*, *Devosia*, *Azorhizobium*, *Microvirga*, *Methylobacterium*, et *Bradyrhizobium* appartiennent aux α-protéobactéries et les deux genres *Burkholderia* et *Cupriavidus* appartiennent aux β-protéobactéries (Remigi *et al.*, 2016 ; Figure 17). Les premières rhizobia découvertes étaient des α-protéobactéries, les β-rhizobia n'ayant été identifiées qu'au début des années 2000 (Moulin *et al.*, 2001).

Les associations symbiotiques entre légumineuses et rhizobia peuvent être plus ou moins spécifiques. Chaque souche de rhizobium peut s'associer avec un nombre limité de plantes hôtes et réciproquement. Il existe des associations symbiotiques très spécifiques ; c'est par exemple le cas *d'Azorhizobium caulinodans* qui ne s'associe qu'avec un certain nombre d'espèces de légumineuses du genre *Sesbania*, dont l'espèce modèle *Sesbania rostrata* (Sprent *et al.*, 2017). Au contraire, les souches *Ensifer fredii* NGR234 et USDA257 ont un spectre d'hôte très large et sont capables de s'associer avec des plantes hôtes appartenant à au moins 112 et 79 genres, respectivement (Pueppke et Broughton, 1999 ; Poole *et al.*, 2018). Du côté de l'hôte, la plante médicinale *Sophora alopecuroides* peut s'associer avec plus de 75 espèces différentes d'α-rhizobia (Zhao *et al.*, 2010). Au sein d'une même espèce de rhizobium,

différentes souches alors nommées biovar (bv.), peuvent s'associer avec des espèces de légumineuses phylogénétiquement éloignées. C'est le cas de l'espèce *Rhizobium leguminosarum* qui possède les trois biovars *phaseoli*, *trifolii* et *viciae* associés respectivement aux genres *Phaseolus*, *Trifolium et Vicia* appartenant à différents clades de légumineuses (Rogel *et al.*, 2011). Chez les rhizobia, l'appartenance à une même espèce validée par la phylogénie de l'ADNr 16S ne corrèle pas du tout avec les phylogénies des gènes *nod* (présentés ci-dessous) et *nifH*. Ces derniers groupent ensemble les différents biovars, c'est-à-dire les souches compatibles avec une même espèce végétale (Rogel *et al.*, 2011). Cette observation peut s'expliquer par la forte mobilité des gènes bactériens *nod* impliqués dans le dialogue moléculaire avec l'hôte et *nif* codant le complexe nitrogénase permettant la fixation d'azote via des transferts horizontaux.

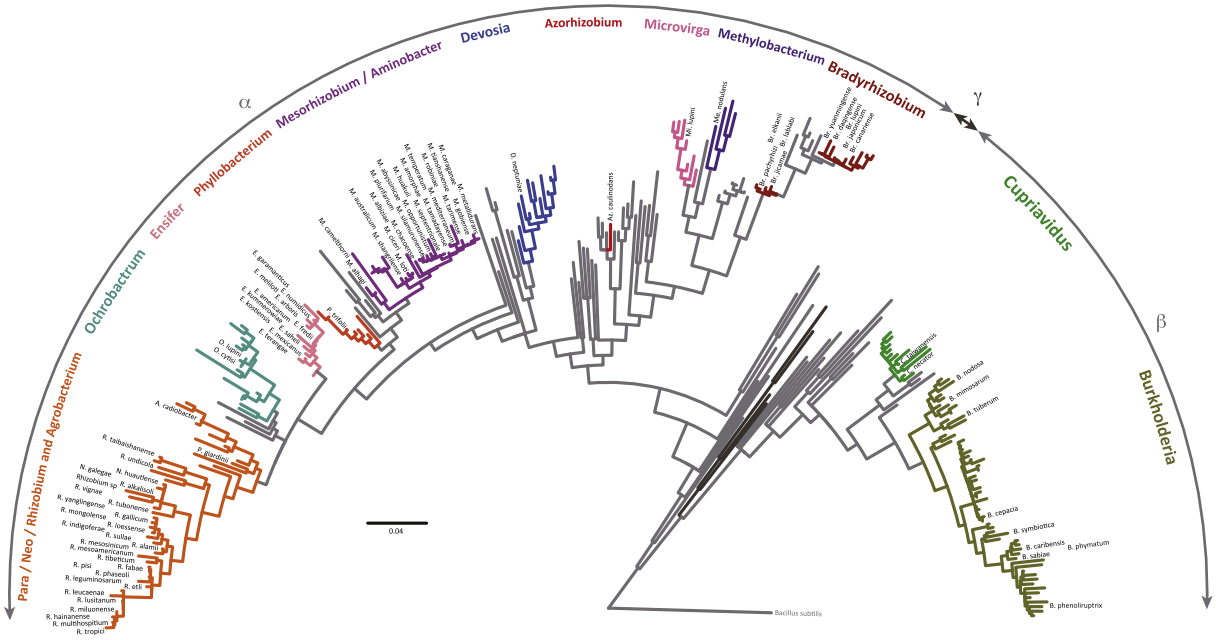


Figure 17: Phylogénie basée sur l'ADNr 16S des Protéobactéries montrant la distribution des rhizobia. Les genres connus de rhizobia sont indiqués en couleur et les espèces associées sont indiquées dans l'arbre (d'après Remigi *et al.*, 2016).

2.3. Le dialogue moléculaire entre les légumineuses et les rhizobia

2.3.1. La libération de flavonoïdes et leur perception par les rhizobia

En cas de carence azotée, les plantes de la famille des légumineuses libèrent un cocktail spécifique de flavonoïdes au sein de leurs exsudats racinaires, qui sont perçus par les rhizobia compatibles de la rhizosphère. Cette perception de flavonoïdes compatibles implique une

interaction avec les facteurs de transcription *nodD1* ou *nodVW* qui activent ensuite une cascade d'expression des gènes de nodulation responsables de la synthèse de molécules de signalisation appelées facteurs Nod (NF), qui à leur tour sont perçus par la plante (Peters *et al.*, 1986 ; Loh et Stacey, 2003 ; Cren *et al.*, 1995 ; Figure 18).

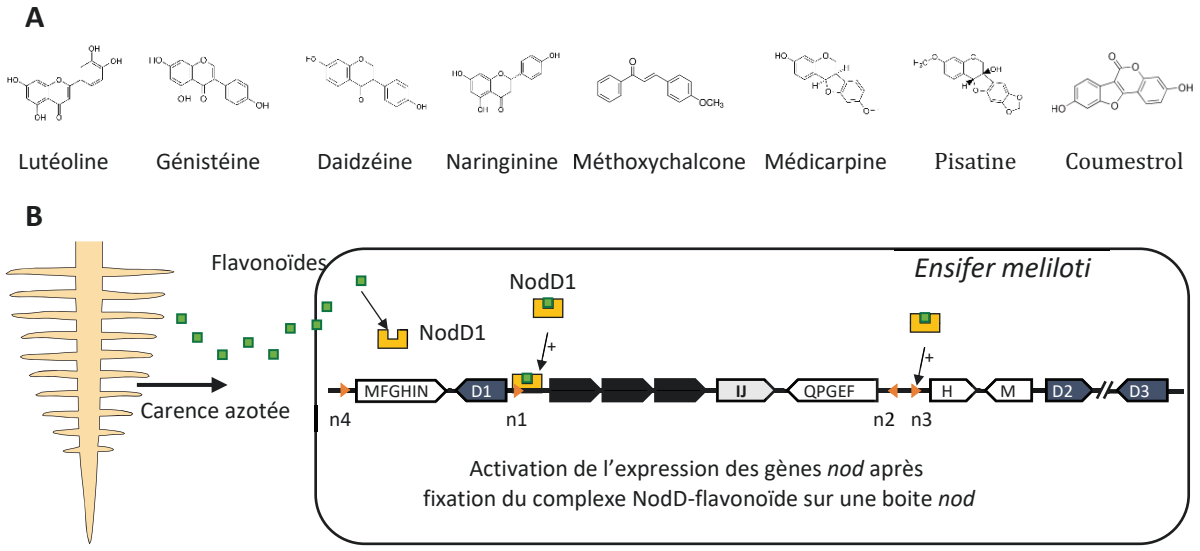


Figure 18: L'induction des gènes de nodulation par les flavonoïdes. (A) Liste de flavonoïdes communs induisant le processus de nodulation chez diverses légumineuses (d'après Liu et Murray., 2016). (B) Schéma simplifié de l'initiation du dialogue moléculaire entre *E. meliloti* et une légumineuse compatible sécrétant des flavonoïdes. Ces derniers interagissent avec le facteur de transcription NodD en formant un complexe activant la transcription des gènes *nod* en se fixant à l'ADN sur les boites nod (pointes de flèches orange). Les gènes *nod* canoniques (Présents chez toutes les bactéries sécrétant des facteurs Nod : *nodABC*, en noir) ou spécifiques (en blanc) sont alors exprimés. Les gènes *nodD* au nombre de trois chez *E. meliloti* sont exprimés de façon concomitante, et les gènes *nodIJ* codent un exporteur des facteurs Nod (d'après Cren *et al.*, 1995).

2.3.2. La sécrétion de facteurs Nod initie l'interaction entre les partenaires

Les facteurs Nod sont composés d'un cœur Lipo-Chito-Oligosaccharidique (LCO ; Figure 19). Ils sont sécrétés par les bactéries et diffusent dans la rhizosphère jusqu'au contact des cellules racinaires de légumineuses (Lerouge *et al.*, 1990). Les protéines NodABC produisent le LCO. NodB et NodC oligomérisent le corps chitinique à partir de N-acétylglucosamine, et NodA permet l'acylation de ce dernier (Mergaert *et al.*, 1997). De nombreux autres gènes *nod* existent et sont impliqués dans les divers motifs de décoration des facteurs Nod, qui sont multiples au sein d'une même souche bactérienne. C'est notamment le cas des protéines NodUSZPQ, NoeEIJKL et NolKLO de la souche *Ensifer fredii* NGR 234 (Perret *et al.*, 2000 ; Broughton *et al.*, 2000 ; Figure 19). Enfin, les protéines NodIJ forment un transporteur responsable de leur sécrétion dans le milieu extracellulaire. La multiplicité des flavonoïdes

exsudées et des facteurs Nod sécrétés sont des éléments fondamentaux de la compatibilité entre une légumineuse et une souche bactérienne donnée (Pueppke et Broughton, 1999).

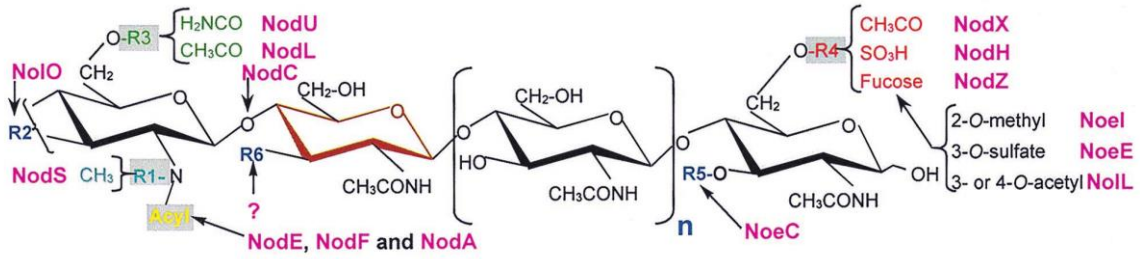


Figure 19: Structure du facteur Nod de *Ensifer fredii* NGR234 et protéines impliquées dans sa synthèse. Les protéines NodABC synthétisent le cœur LCO, et les autres gènes présentés ici sont impliqués dans les décorations potentielles des facteurs Nod (carbonylation, glycosylation, fucosylation, méthylation) (d'après Perret *et al.*, 2000).

Les facteurs Nod sont perçus par la plante via des récepteurs de type LysM-RLK appelés NFP/LYK3 (*Medicago truncatula*) ou NFR1/NFR5 (*Lotus japonicus*), et impliquent également le récepteur LRR-RLK DMI2 (Limpens *et al.*, 2005 ; Oldroyd, 2013). Cette reconnaissance active une cascade de signalisation qui débute par des oscillations calciques dans le noyau impliquant un grand nombre de protéines membranaires telles que DMI1 (Ané *et al.*, 2004), CNGC15 (Charpentier *et al.*, 2016) et MCA8 (Capoen *et al.*, 2011 ; Zipfel et Oldroyd, 2017 ; Figure 20). Il a été suggéré que la transduction du message des facteurs Nod de la membrane plasmique jusqu'au noyau se fait par l'intermédiaire du mévalonate (Venkateshwaran *et al.*, 2015). En effet HMGR1 (3-Hydroxy-3-Methylglutaryl CoA Reductase 1), qui est une enzyme appartenant à la voie de biosynthèse du mévalonate, interagit directement avec DMI2 et l'ajout exogène de mévalonate est un signal suffisant pour induire les oscillations calciques au niveau de la membrane nucléaire via DMI1 (Venkateshwaran *et al.*, 2015 ; Figure 20). Ces signaux calciques nucléaires sont décodés par une protéine kinase appelée CCAMK (Lévy *et al.*, 2004). Dans le cas de la symbiose fixatrice d'azote, CCAMK phosphoryle le facteur CYCLOPS qui recrute alors les facteurs de transcription à domaine GRAS appelés DELLA, NSP1 et NSP2 (Zipfel et Oldroyd, 2017). Ces derniers permettent l'expression de gènes symbiotiques tels que NIN et ERN1 (Zipfel et Oldroyd, 2017 ; Poole *et al.*, 2018). Cette signalisation précoce enclenche deux mécanismes coordonnés : l'infection des tissus végétaux et l'organogénèse de la nodosité. L'organogénèse de la nodosité implique la dédifférenciation et la division de cellules racinaires proches de la zone corticale. Rapidement, la mise en place d'une nouvelle assise cellulaire mène à terme à la formation de la zone méristématique de la nodosité.

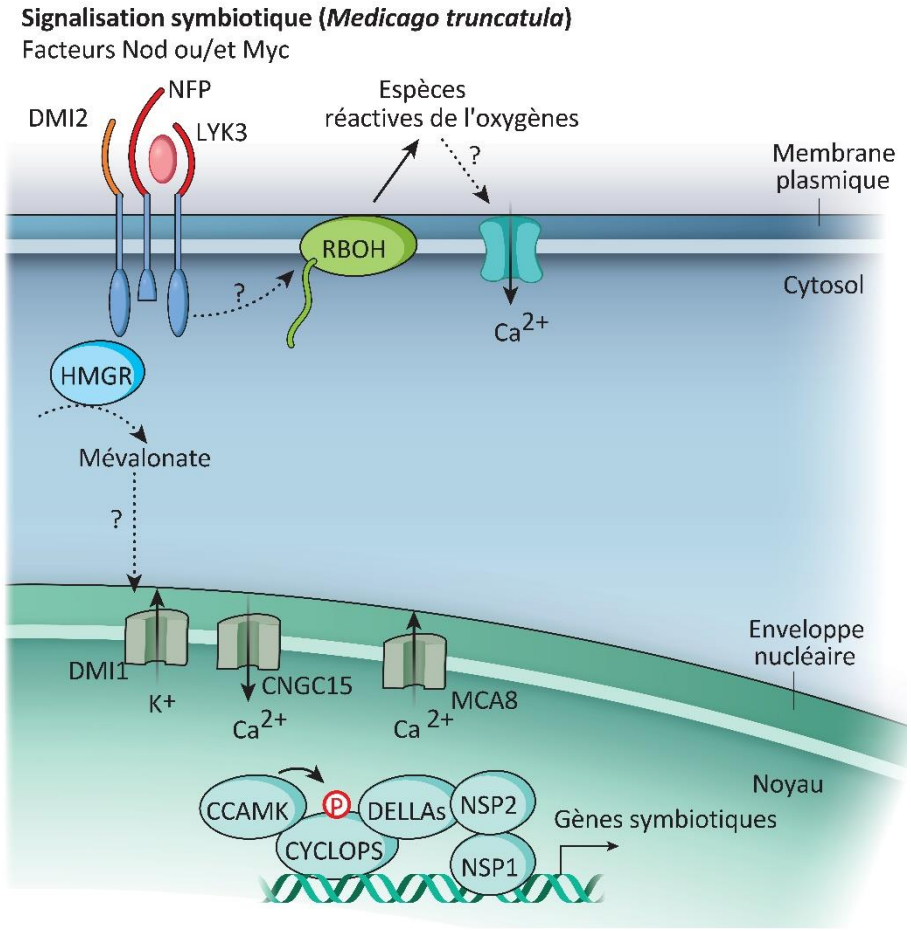


Figure 20: Schéma simplifié de la réponse à la perception des facteurs Nod chez *Medicago truncatula* (d'après Zipfel et Oldroyd, 2017).

2.3.3. La voie de signalisation commune des endosymbioses

En dehors du processus de nodulation, les légumineuses sont également capables, comme 85% des espèces végétales terrestres d'effectuer des symbioses dites mycorhiziennes avec des Eumycètes. Il existe divers degrés d'intimité entre les végétaux et leurs symbiontes mycorhiziens. Les ectomycorhizes forment un manteau fongique autour des racines et leurs hyphes peuvent pénétrer les espaces intercellulaires situés autour des cellules corticales racinaires. Les symbioses ectomycorhiziennes sont apparues assez récemment de façon convergente de nombreuses fois il y a environ 50 Ma (Strullu-Derrien *et al.*, 2018). Ces symbioses ectomycorhiziennes sont cependant en majorité cantonnées aux régions tempérées du globe et restreintes à des espèces ligneuses (gymnospermes et angiospermes). Les mécanismes de leur mise en place est peu connu. Les symbioses mycorhiziennes les plus communes et les plus étudiées sont les endomycorhizes arbusculaires (ou à arbuscules) dont les

symbiontes font partie de la famille des Glomérales. Les hyphes se développent dans l'espace apoplastique et pénètrent les cellules végétales en formant des arbuscules, structures ramifiées et surfaces d'échange entre la plante et le champignon. Comparée à la nodulation dont l'origine remonte à ~70 Ma, la symbiose endomycorhizienne serait très ancienne. Elle a d'ailleurs été observée dans des fossiles de la flore de Rhynie (Ecosse) datant de plus de 400 Ma (Strullu-Derrien *et al.*, 2018).

De façon remarquable, le dialogue moléculaire et la signalisation de l'interaction entre les végétaux et les champignons endomycorhiziens (e.g *Rhizophagus irregularis*) sont en grande partie similaires à ceux de la symbiose entre rhizobia et légumineuses (Oldroyd, 2013). La voie de signalisation de la symbiose rhizobium-légumineuses a donc emprunté le « châssis » moléculaire établi pour la mise en place de l'endomycorhization. Les champignons endomycorhiziens perçoivent des strigolactones présentes dans les exsudats racinaires qui induisent la sécrétion des facteurs Myc (Maillet *et al.*, 2011). Ces facteurs Myc, qui sont un mélange de LCO et de Chito-Oligosaccharides non sulfatés, sont perçus à leur tour par d'autres récepteurs de type LysM (Figure 20). Ces derniers activent également des oscillations calciques dans le noyau nécessitant les mêmes transporteurs DMI1, CNGC15 et MCA8. Les oscillations sont perçues par CCamK phosphorylant CYCLOPS, qui recrute une combinaison de facteurs de transcription initiant la réponse cellulaire différente de celle de la nodulation, impliquant les protéines NSP1 et RAM1 (Oldroyd, 2013).

2.3.4. La formation des cordons d'infection et morphogénèse des nodosités

Le contact des rhizobia compatibles avec l'hôte végétal au niveau des poils absorbants induit une courbure de ces derniers en forme de crosse de berger (Figure 21). Les bactéries sécrètent une cellulase et accèdent à la membrane plasmique du poil absorbant qui s'invagine, faisant ainsi entrer les bactéries vers les tissus plus profonds (Robledo *et al.*, 2008). Cette invagination membranaire est nommée cordon d'infection. Ce dernier va croître et se ramifier sous l'action combinée des deux partenaires symbiotiques. En effet, une pectate lyase impliquée dans la progression des cordons d'infection a été découverte chez *Lotus japonicus*, suggérant donc également un rôle actif de la plante dans le remodelage de sa paroi au cours du processus infectieux (Xie *et al.*, 2012). Une population de rhizobia va s'y diviser et rester au contact des extrémités des invaginations, qui pénètrent progressivement à l'intérieur des tissus racinaires.

De nombreux gènes des deux partenaires nécessaires à plusieurs étapes clés du processus d'infection sont connus (Jones *et al.*, 2007 ; Figure 21A). C'est notamment le cas des gènes bactériens *exoY* et *exoH*, impliqués dans la synthèse des exopolysaccharides de surface comme les succinoglycanes. Ces derniers limitent l'activation de réactions immunitaires de l'hôte végétal. Du côté de ce dernier, le récepteur d'exopolysaccharides bactériens EPR3 a été identifié chez *L. japonicus*. Il est requis pour une symbiose fonctionnelle et perçoit les bactéries compatibles ou non avec l'hôte, inhibant la croissance du cordon d'infection le cas échant (Kawaharada *et al.*, 2015). L'élongation du cordon d'infection nécessite une coordination de la biosynthèse de composés pariétaux, lipidiques et cytoplasmiques apportés par des vésicules transitant via le cytosquelette d'actine et de microtubules (Oldroyd *et al.*, 2011 ; Figure 21B).

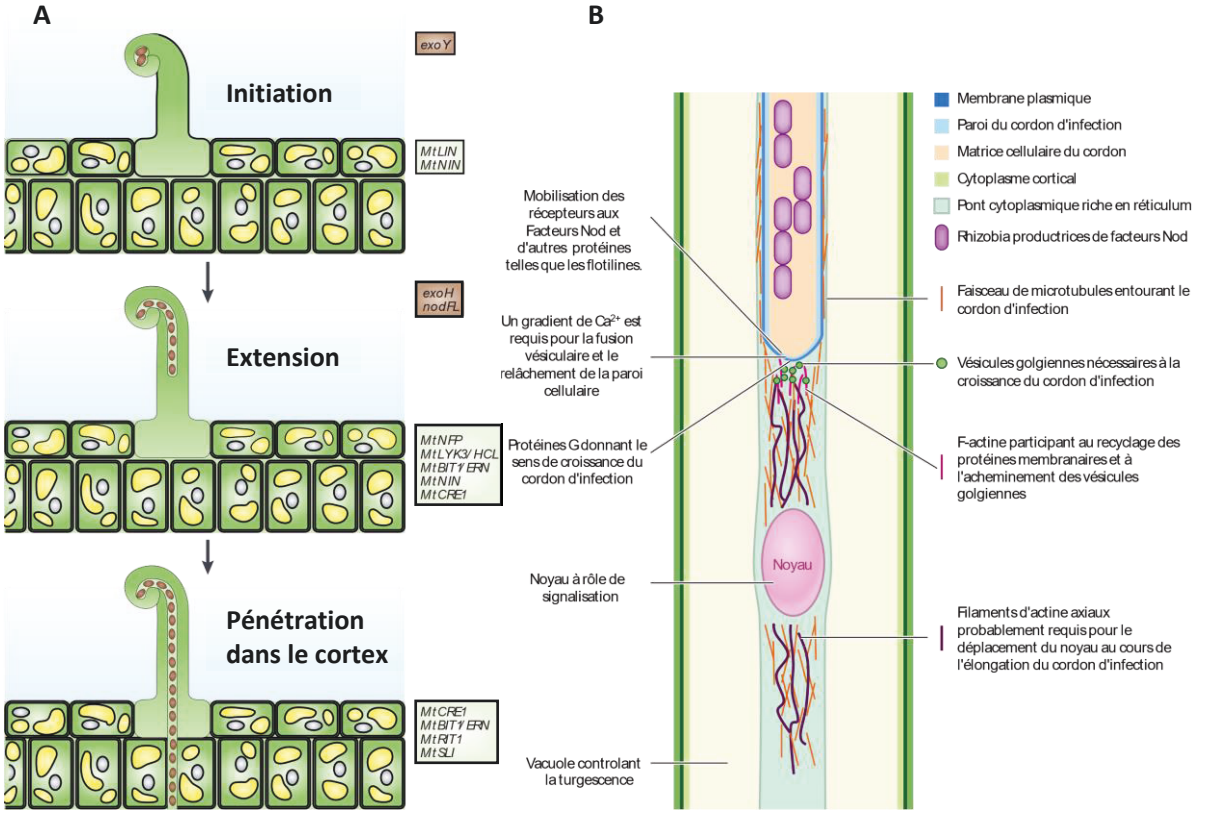


Figure 21: Infection des poils racinaires de *Medicago truncatula* par *Ensifer meliloti*. (A) Gènes nécessaires aux différentes étapes de la formation du cordon d'infection. Les gènes dans les encadrés verts et marrons sont respectivement ceux de l'hôte végétal et du symbionte bactérien (d'après Jones *et al.*, 2007). (B) Modèle de croissance du cordon d'infection (d'après Oldroyd *et al.*, 2011).

Les bactéries infectent les cellules symbiotiques de façon intracellulaire en passant par les cordons d'infection et deviennent alors des bactéroïdes (Figure 22B). Ce processus proche de l'endocytose implique le gain d'une membrane d'origine eucaryote pour former le symbiosome

composé du bactéroïde et de sa membrane du symbiosome. L'espace entre bactéroïde et la membrane du symbiosome est nommé espace péri-bactéroïdien. Cette infection requiert des protéines vésiculaires de type SNARE telles que MtVAMP72 (Ivanov *et al.*, 2012). En dehors de l'infection via les poils absorbants, il existe d'autres modes de pénétration des bactéries. Le processus du *crack-entry* existe notamment chez *Sesbania rostrata* et chez les nodosités de type Aeschynomenoïdes, où les points d'infection se situent au niveau des blessures du rhizoderme dans les zones d'émergence des racines secondaires, menant à la formation de nodosités visibles à l'aisselle de ces dernières (Deakin et Broughton., 2009, Figure 22A).

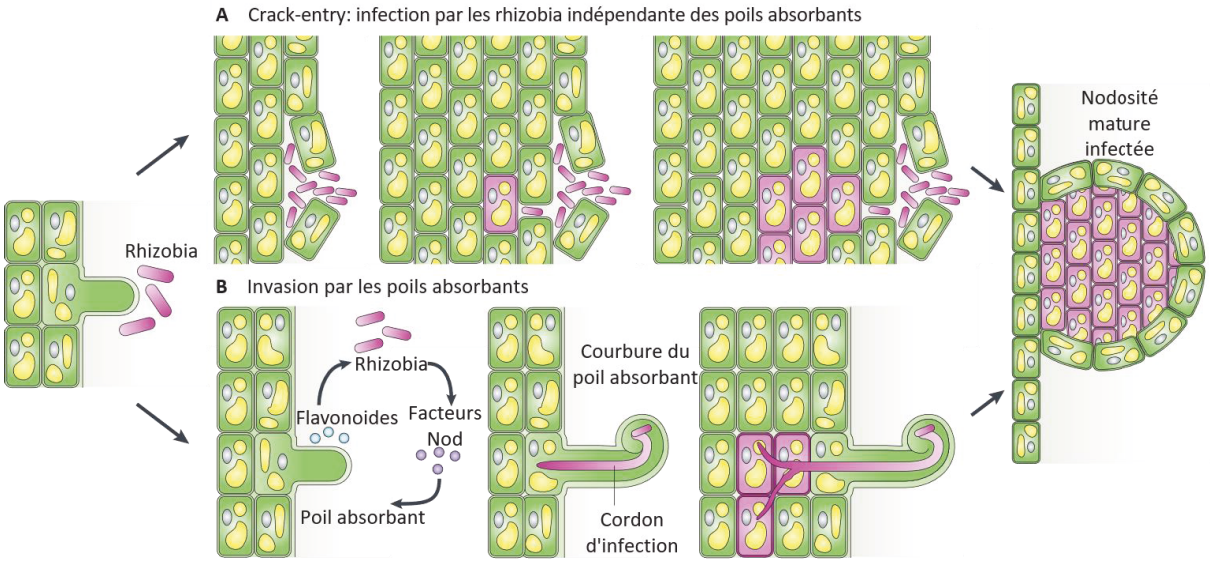


Figure 22: Différents modes d'infection des racines par les rhizobia parallèles à l'organogénèse de nodosités fixatrices d'azote. (A) Infection par *crack-entry*. (B) infection par formation d'un cordon d'infection initié au niveau d'un poil absorbant racinaire (d'après Deakin et Broughton., 2009).

En fonction de l'espèce végétale, les nodosités peuvent être dites indéterminées (e.g. *Medicago truncatula*, *Pisum sativum*) ou déterminées (e.g. *Lotus japonicus*, *Glycine max*, *Phaseolus vulgaris*, *Aeschynomene* spp.). La distinction est basée sur l'existence d'un méristème apical persistant chez la nodosité indéterminée mature et fixatrice d'azote. La nodosité indéterminée se découpe en plusieurs zones visibles en coupe longitudinale (Kondorosi *et al.*, 2013 ; Figure 23B). De l'apex de la nodosité jusqu'à sa région la plus distale, on peut observer 5 zones. La première est constituée par un méristème apical (I), sans bactéries. Ensuite se trouve la zone d'infection (II), où les cellules sont parcourues par les cordons d'infection relarguant les bactéries, qui se différencient ensuite en bactéroïdes. La troisième zone est celle où les bactéroïdes sont matures et la fixation d'azote a lieu (III). La zone IV est

dite de sénescence dans les nodosités plus âgées, où les bactéroïdes sont progressivement digérés par les cellules végétales de façon radiale, du centre vers la périphérie (Dupont *et al.*, 2012). Une autre zone a été mentionnée comme étant la zone saprophytique (V), où la sénescence est terminée, et qui est remplie de rhizobia saprophytes non fixatrices proliférant dans ces tissus végétaux en décomposition (Timmers *et al.*, 2000). Les nodosités déterminées présentent un développement synchrone des cellules symbiotiques et ne conservent pas leur méristème nodositaires une fois matures (Kondorosi *et al.*, 2013). Les deux types de nodosités sont retrouvées dans tous les biomes, bien qu'on observe une prédominance des nodosités indéterminées en régions tempérées, et déterminées en régions subtropicales (Kondorosi *et al.*, 2013). Après l'infection, les cellules symbiotiques se différencient en même temps que les bactéroïdes se divisent dans les symbiosomes.

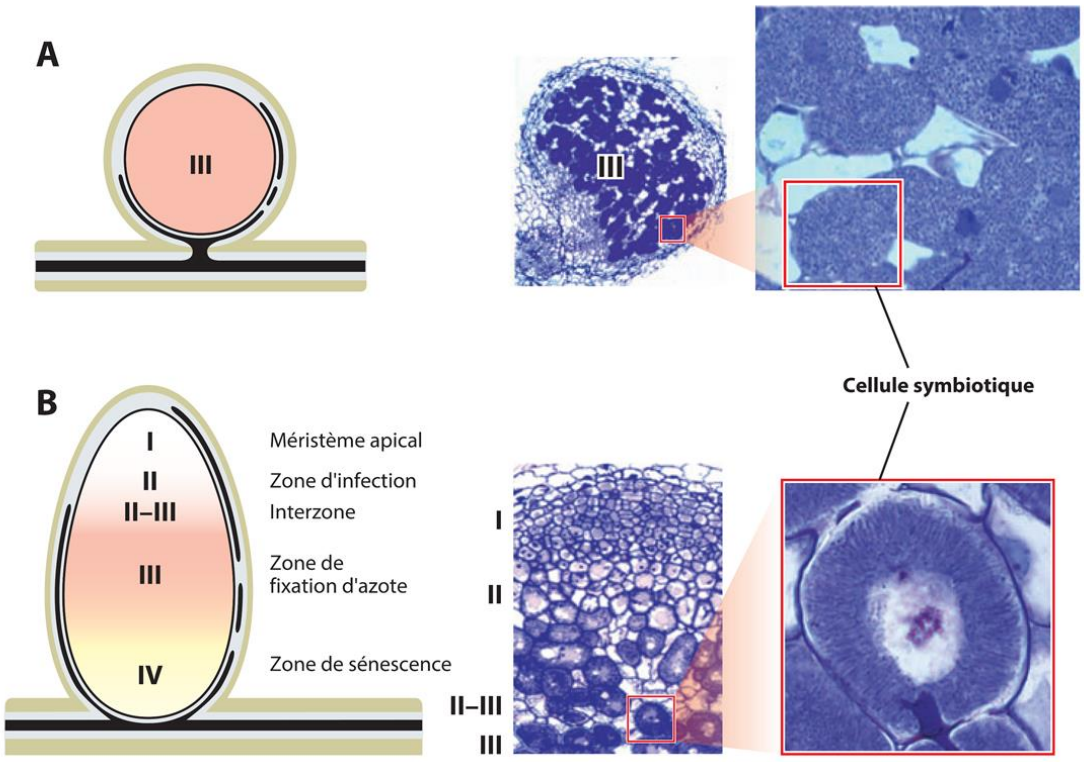


Figure 23: Structure des nodosités racinaires déterminées et indéterminées. (A) Nodosité déterminée sphérique dont la région centrale (III) contient des cellules non infectées et des cellules symbiotiques fixatrices d'azote car infectées par des rhizobia, visibles sur l'agrandissement. (B) Les nodosités indéterminées ont un méristème apical (I), une zone d'infection (II), une interzone (II-III), une zone de fixation d'azote (III) et une zone de sénescence dans les nodosités plus âgées (IV). Chez *Medicago*, les volumes cellulaires et nucléaires des cellules symbiotiques infectées augmentent progressivement car ces dernières subissent des cycles d'endoréplication. Les faisceaux conducteurs à la périphérie des nodosités sont représentés en noir (d'après Kondorosi *et al.*, 2013).

2.3.5. La différenciation des cellules symbiotiques

Chez de nombreux organismes eucaryotes, il est fréquent d'observer des cellules différenciées polyploïdes comme les cellules musculaires striées humaines, ou encore les cellules du fruit de la tomate au cours de sa maturation. Les cellules diploïdes du futur péricarpe sont d'abord non différenciées et ont une quantité d'ADN de 2C ou 4C selon le stade du cycle cellulaire dans lequel elles se situent (ie. avant ou après la phase S de réplication de l'ADN). La valeur 1C constitue la quantité d'ADN totale du génome haploïde. Les cellules du péricarpe de tomate entrent en phase d'expansion cellulaire sans division, menant à des cellules dont la ploïdie atteint 256C quand le fruit est mature (Joubès et Chevalier, 2000).

Dans le contexte de la symbiose rhizobium-légumineuses, la différenciation des cellules symbiotiques se manifeste par une augmentation de leur taille et de leur niveau de ploïdie (Cebolla *et al.*, 1999 ; González-Sama *et al.*, 2006 ; Figure 24A). Une protéine activatrice du complexe promoteur de l'anaphase nommée Ccs52A a été découverte chez *Medicago sativa* et *Medicago truncatula* et est impliquée dans la transition d'un cycle cellulaire classique à des endocycles (Cebolla *et al.*, 1999 ; Vinardell *et al.*, 2003).

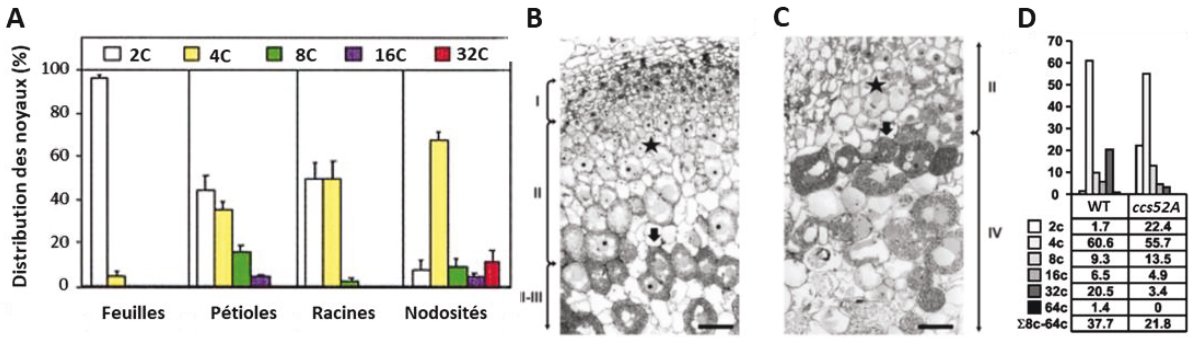


Figure 24: Les cellules symbiotiques des nodosités de *M. truncatula* subissent des cycles d'endoréplication requise pour une symbiose fonctionnelle, et cette transition est permise par la protéine Ccs52A. (A) Index de ploïdie des noyaux de *M. truncatula* dans divers organes, montrant un fort taux d'endoréplication dans les nodosités allant jusqu'à 64C. (B, C, D) Comparaison fonctionnelle entre les nodosités de *M. truncatula* sauvage et transformée avec l'ARN antisens de *ccs52A*. (B, C) Comparaison histologique entre les conditions sauvage et RNAi-Ccs52A montrant l'existence d'une zone de sénescence précoce chez la plante mutante, ainsi que l'absence de zone III de fixation d'azote. Les chiffres romains présentent les zones des nodosités indéterminées présentées précédemment. (D) Index d'endoréplication montrant une chute du nombre de noyaux endorépliqués chez la plante mutante (d'après Cebolla *et al.*, 1999 ; Vinardell *et al.*, 2003).

Ce phénomène a également été découvert dans les nodosités déterminées de *Lotus japonicus* (González-Sama *et al.*, 2006). La protéine Ccs52A est très exprimée dans les nodosités mais

n'est pas requise pour leur organogénèse. Elle est cependant nécessaire à la différenciation des cellules symbiotiques et à leur fonctionnement. En effet, l'expression d'un ARN antisens de *ccs52A* induit une sénescence précoce des nodosités de *M. truncatula*, via l'absence de zone III observable, ainsi qu'un niveau d'endoréplication plus faible que la condition sauvage (Vinardell *et al.*, 2003 ; Figure 24B-D).

2.3.6. L'expression des gènes bactériens impliqués dans la fixation d'azote

Une fois les bactéries endocytées, les bactéroïdes commencent à opérer leur spécialisation métabolique orientée principalement vers la fixation de l'azote atmosphérique. La concentration en O_2 doit être réduite dans les nodosités car il inactive le complexe nitrogénase. Cependant, les bactéroïdes nécessitent tout de même du dioxygène pour produire de l'ATP requis pour la fixation d'azote. Ce « paradoxe oxygénique » impose ainsi un intervalle de concentrations d'O_2 permissives dans les nodosités des légumineuses, oscillant entre 5 et 30nM (Marchal et Vanderleyden, 2000 ; Terpolilli *et al.*, 2012). Ces concentrations sont obtenues grâce à la complexe barrière de diffusion de l'oxygène du cortex de la nodosité (Minchin *et al.*, 2008) et diverses protéines provenant des deux partenaires symbiotiques.

Du côté végétal, les leghémoglobines précédemment présentées ont une forte affinité pour le dioxygène, limitant alors le niveau d'oxygène libre dans les nodosités. Du côté des bactéroïdes, l'O_2 est très rapidement consommé via l'expression un complexe oxydase à haute affinité pour l'O_2 (*fixNOPQ*) (Delgado *et al.*, 1998). Il a été découvert que la régulation de l'expression des gènes *nif* et *fix* permettant la formation du complexe nitrogénase était sous la dépendance d'une faible pression partielle en O_2. Il existe tout de même une certaine diversité de mécanismes de détection du niveau d'oxygène parmi les rhizobia, principalement au niveau des cascades de régulation (Terpolilli *et al.*, 2012). De façon générale, FixL est le premier senseur d'O_2 et active FixJ dont les cibles varient ensuite en fonction de l'espèce de rhizobia concernée. Chez *E. meliloti*, FixJ active à la fois l'expression de FixK qui induit l'expression de l'oxydase terminale *fixNOPQ* et de NifA qui induit à son tour l'expression des gènes *nif* (Figure 25A). Chez *B. diazoefficiens*, FixJ active l'expression en cascade de FixK2 et de FixK1 qui activent l'expression de *fixNOPQ* et contrôlent négativement l'expression des gènes *nif*, respectivement. L'expression de NifA est ici sous contrôle du système à deux composantes RegSR, et la stabilité de NifA est dépendante de la pression partielle en dioxygène (Figure

25B). Concomitamment à l'expression des gènes responsables de la fixation d'azote, il existe une mise en place de nombreux échanges de métabolites entre les cellules symbiotiques végétales et leurs bactéroïdes endocytés.

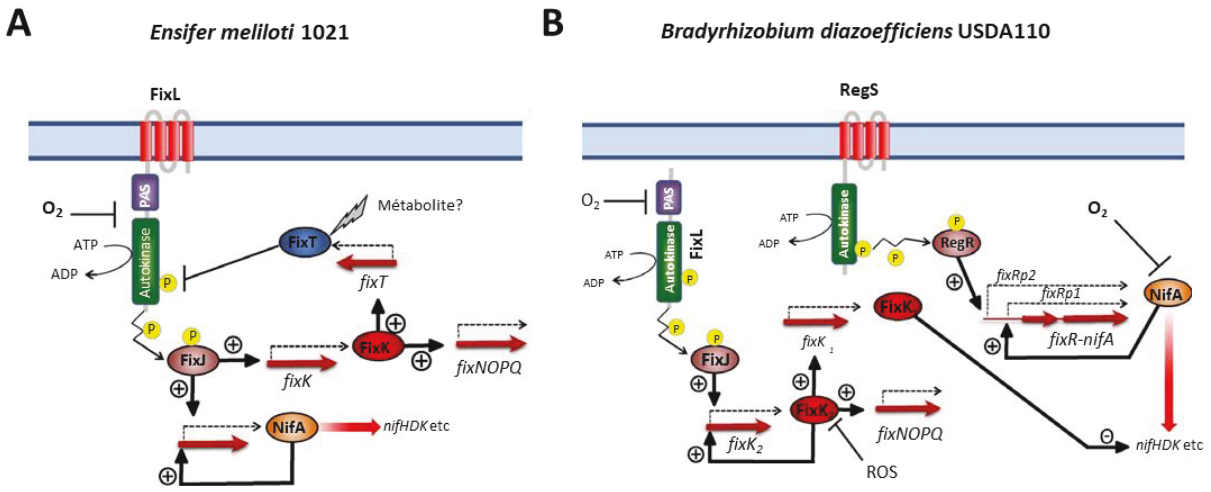


Figure 25: Comparaison des cascades de régulation de l'expression des gènes de fixation d'azote des rhizobia entre *E. meliloti* 1021 (A) et *B. diazoefficiens* USDA110 (B). FixL est le premier senseur d'O_2, et il est transmembranaire chez *E. meliloti* mais soluble chez *B. diazoefficiens*. Il phosphoryle FixJ qui active FixK chez *E. meliloti*, et FixK2 qui active lui-même FixK1 chez *B. diazoefficiens*. FixK ou FixK2 induisent l'expression des gènes *fixNOPQ* codant une oxydase terminale à haute affinité pour l'O_2. FixJ active également l'expression du régulateur majeur NifA des gènes *nif* chez *E. meliloti*, là où la régulation de l'expression de NifA est plus complexe chez *B. diazoefficiens* car contrôlée par le système à deux composantes RegSR. De plus, l'expression des gènes *nif* est aussi négativement controlée par FixK1. Chez *B. diazoefficiens*, NifA est sensible à au dioxygène (d'après Terpolilli *et al.*, 2012).

2.3.7. Les échanges nutritionnels hôte-symbionte au niveau des symbiosomes

La majorité des rhizobia ne sont pas diazotrophes au sens propre. Elles ne peuvent pas fixer librement l'azote car ne possèdent pas le gène *nifV* codant une homocitrate synthase, dont le produit (R-homocitrate) est requis pour la formation du Fe-Mo-Co. Il existe quelques exceptions, notamment chez les bradyrhizobia photosynthétiques capables de fixation libre (Alazard, 1990 ; Nouwen *et al.*, 2017). La fixation d'azote a tout de même lieu dans les nodosités grâce à un apport d'homocitrate végétal produit via le gène *FEN1* qui a été découvert chez *L. japonicus* (Hakoyama *et al.*, 2009). La souche photosynthétique *Bradyrhizobium* sp. ORS285 possède le gène *nifV* mais un mutant Δ*nifV* peut tout de même fixer l'azote si son hôte possède également le gène FEN1, comme c'est le cas pour l'hôte *Aeschynomene afraspera* (Nouwen *et al.*, 2017). Différents ions requis pour l'activité nitrogénase ou les fonctions de ménage des bactéroïdes sont obtenus par des transporteurs de l'hôte vers l'espace péri-

bactéroïdien puis importés via d'autres transporteurs bactériens dans le cytoplasme du bactéroïde. Les sources de carbone des bactéroïdes sont en général des dicarboxylates (ex : malate) transportés par le transporteur bactérien DctA. Malgré la production d'azote assimilable, les bacteroïdes ne synthétisent plus certains de leurs propres acides aminés et certains semblent être fournis par l'hôte. C'est le cas des acides aminés branchés (Ile, Leu, Val) requis pour une symbiose fonctionnelle chez le pois (Prell *et al.*, 2009 ; Udvardi et Poole, 2013 ; Figure 26). Ce phénomène s'appelle l'auxotrophie symbiotique. Elle résulte d'une spécialisation du métabolisme des bactéroïdes qui passent d'une grande diversité de fonctions activées en vie libre à une hyperfocalisation métabolique lors de fixation d'azote. Le cycle de Krebs n'a alors plus qu'un rôle catabolique. Au sein des nodosités fonctionnelles, les bactéroïdes fournissent de l'azote assimilable à la plante sous forme de NH_4^+ qui peut être exporté sous forme de glutamine, d'asparagine ou d'uréides par le système vasculaire (Oldroyd *et al.*, 2011 ; Udvardi et Poole, 2013 ; Figure 26).

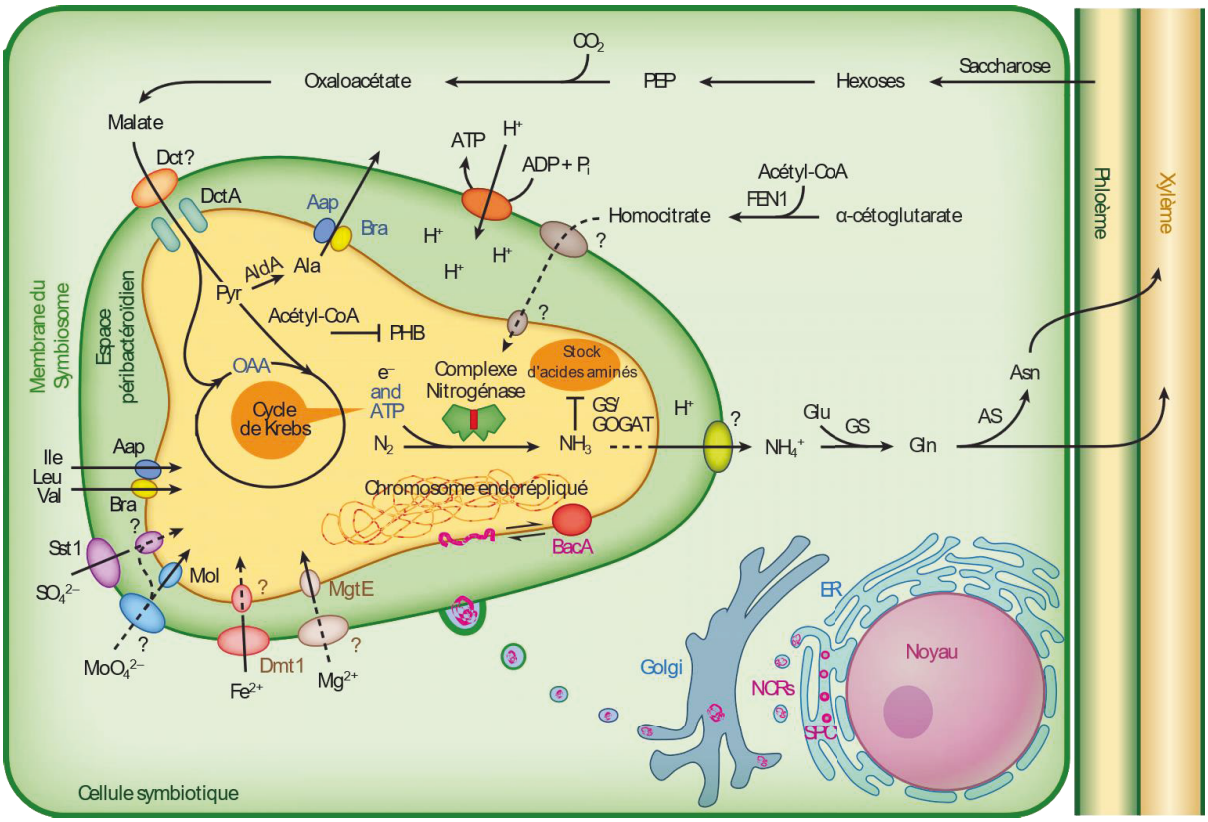


Figure 26: Schéma récapitulatif des échanges nutritionnels intervenants au niveau du symbiosome entre bactéroïde et cellule symbiotique. Les protéines et métabolites les plus importants et caractérisés comme nécessaires au processus symbiotique sont présentés (d'après Oldroyd *et al.*, 2011).

3. Les processus de différenciation cellulaire des cellules symbiotiques et des bactéroïdes

3.1. La diversité des morphotypes de bactéroïdes

Au sein des nodosités de légumineuses, les bactéroïdes peuvent adopter différentes morphologies (morphotypes) en fonction de l'interaction hôte-symbionte. Trois morphotypes ont été observés jusqu'à présent dans la nature (Oono *et al.*, 2010). Il existe des bactéroïdes semblables aux bactéries en culture (ou dans le sol), dits indifférenciRés (ou U pour « Unmodified »), comme c'est le cas de *Mesorhizobium loti* et *Bradyrhizobium diazoefficiens* USDA110 qui s'associent respectivement avec *Lotus japonicus* et *Glycine max* (Oono *et al.*, 2010, Figure 27A). En revanche, les bactéroïdes des légumineuses de type IRLC (voir section 3.3) ne présentent pas les mêmes caractéristiques morphologiques que leurs homologues bactériens en vie libre. C'est notamment le cas des bactéroïdes de *Vicia sativa* (Beijerinck, 1888 ; Oke et Long 1999 ; Figure 27B). Ces bactéroïdes possèdent un morphotype appelé E (*Elongated*). Enfin, certaines plantes de la famille des Dalbergoïdes possèdent dans leurs nodosités des bactéroïdes qui présentent un morphotype sphérique ou S (*Spherical*) comme *Arachis hypogea* ou *Aeschynomene indica* (Oono *et al.*, 2010, Bonaldi *et al.*, 2011, Figure 27C). L'hôte végétal impose le programme de différenciation à son symbionte, car une même espèce bactérienne peut adopter, selon l'hôte végétal, des morphotypes différents. Par exemple, la souche *Bradyrhizobium* 32H1 adopte les morphotypes S et U en association avec *Arachis hypogea* et *Vigna unguiculata*, respectivement (Sen et Weaver, 1984). La diversité des morphotypes bactériens et des types de nodosités est résumée ci dessous (Oono *et al.*, 2010 ; Figure 28).

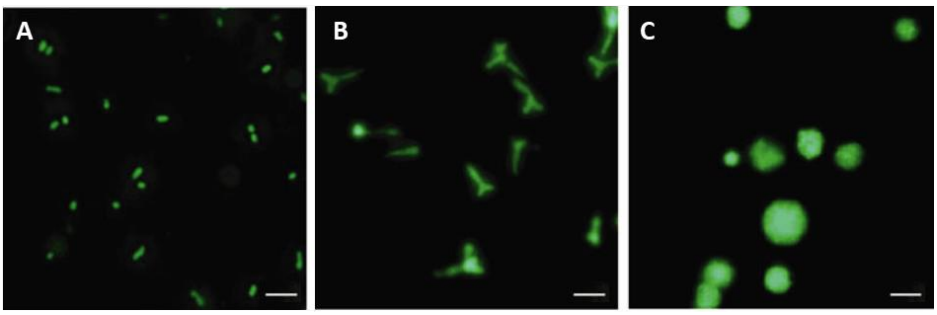


Figure 27: Diversité des morphotypes de bactéroïdes chez les légumineuses. (A) Bactéroïdes non-différenciés de nodosités de *Robinia Pseudoacacia*. (B) Bactéroïdes allongés des nodosités de *Pisum sativum*. (C) Bactéroïdes sphériques isolés des nodosités d'*Arachis hypogea*. Les bactéroïdes sont marqués par le SYTO13, un agent intercalant de l'ADN fluorescent. Barre d'échelle : 5µm (d'après Oono *et al.*, 2010)

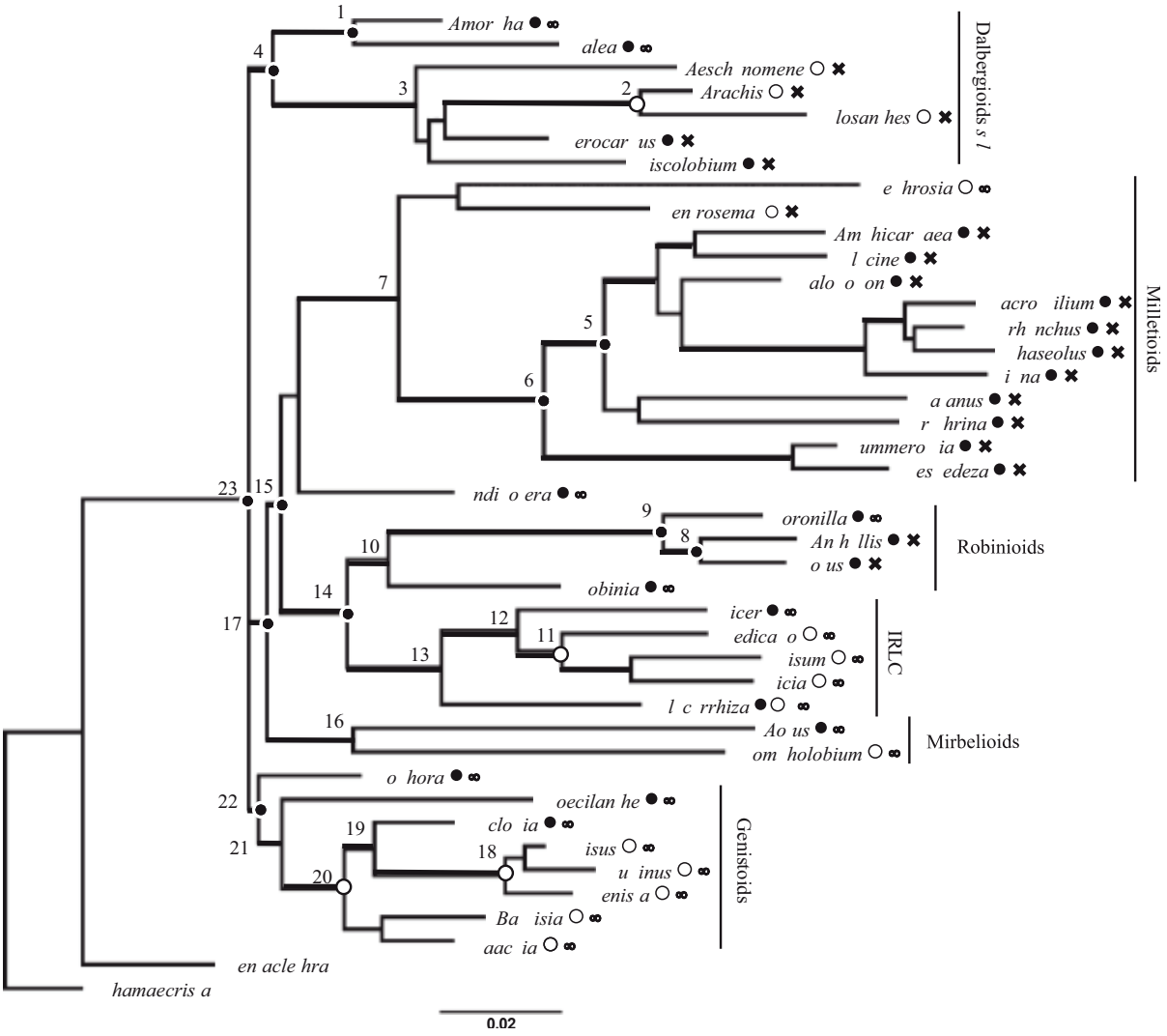


Figure 28: Phylogénie de 40 espèces de *Papilionoideae* résumant le type de nodosité et l'occurrence du processus de différenciation des bactéroïdes. Phylogénie basée sur la combinaison des séquences *matK*, *rbcL* 5.8S rRNA et *trnL*. Les types de nodosités sont représentées avec le symbole ∞ pour les nodosités indéterminées et le symbole × pour les nodosités déterminées. Les plantes qui induisent un morphotype différencié (E ou S) ou indifférencié (U) sont respectivement indiquées par un cercle blanc et noir (d'après Oono *et al.*, 2010).

3.3. La différenciation des bactéroïdes

3.3.1. Observations cytologiques

Le processus de différenciation des bactéroïdes a été d'abord caractérisé chez les endosymbiontes des IRLC. Ces bactéroïdes présentent un allongement cellulaire progressif du haut de la zone d'infection (II) jusqu'à la fin de la zone de fixation (III) (Vasse *et al.*, 1990 ; Figure 29). Il est possible d'observer 5 stades différents de différenciation des bactéroïdes le long du gradient de différenciation spatiale de la nodosité indéteminée de *Medicago sativa* par microscopie électronique en transmission (Vasse *et al.*, 1990 ; Table 2 ; Figure 29). On y voit

un allongement progressif des bactéroïdes de type 1 à type 4, et les bactéroïdes de type 5 semblent vides car présentent une très faible densité aux électrons.

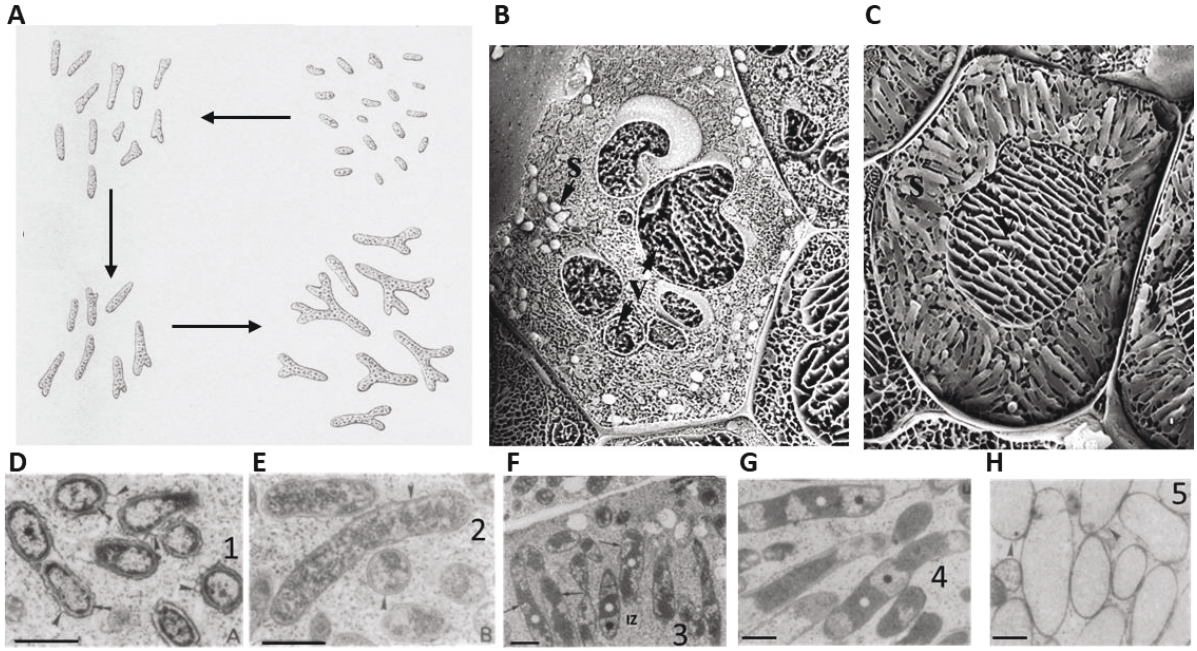


Figure 29: Le processus de différenciation des bactéroïdes dans les nodosités d'IRLC (Inverted Repeat Lacking Clade). (A) Première représentation de la différenciation des bactéroïdes dans les nodosités de *Vicia sativa* faite par Beijerinck en 1888. Les flèches représentent les différents stades de maturation des bactéroïdes commençant à la taille classique de bactérie jusqu'aux bactéroïdes différenciés et branchés en forme de Y (d'après Oke et Long, 1999). (B, C) Photographies de cellules symbiotiques de *Medicago truncatula* observées en microscopie électronique à balayage en cryofracture. (B) Jeunes cellules infectées montrant quelques symbiosomes (s) et plusieurs vacuoles (v). (C) Cellules symbiotiques matures infectées par des bactéroïdes fixateurs d'azote allongés (d'après Coba de la Peña *et al.*, 2018). (D, E, F, G, H) Photographie de microscopie électronique en transmission présentant les différents types de bactéroïdes mentionnés dans la Table 2, respectivement de type 1 à 5. Seuls les bactéroïdes de type 1 se divisent encore au sein des symbiosomes. Barres d'échelle pour les figures D, E, F, G et H : 1µm (d'après Vasse *et al.*, 1990).

Zone de la nodosité	Description de la zone	Stade de différenciation des bactéroïdes	Caractéristiques spécifiques des bactéroïdes dans les différentes zones
I	Méristème	-	Pas ou peu de bactéries
II	Zone d'infection	Non différenciées	Bactéries non différenciées dans les cordons d'infection
		Bactéroïdes de type 1	Tout juste libérés des cordons d'infection
		Bactéroïdes de type 2	Pas de division, en cours d'élongation
II-III	Interzone	Bactéroïdes de type 3	Complètement allongés, hétérogénéité cytoplasmique
III	Zone de fixation	Bactéroïdes de type 4	Allongés, distinctions entre zones riches (ribosomes) et pauvres en électrons
IV	Zone de sénescence	Bactéroïdes de type 5	Bactéroïdes sénescents, faible densité électronique
V	Zone saprophytique	Non différenciées	Bactéries provenant des résidus du cordon d'infection se développent de façon saprophytique sur les tissus en cours de sénescence

Table 2: Répartition spatiale des différents types de bactéroïdes le long d'une nodosité indéterminée et leurs caractéristiques (d'après Lang et Long, 2015).

L'allongement des bactéroïdes est également couplé à une augmentation du niveau de ploïdie, pouvant atteindre 24 copies d'ADN par cellule (Mergaert *et al.*, 2006 ; Figure 30). De plus, cette différenciation est considérée comme terminale dans la mesure où les bactéroïdes différenciés semblent incapables de reprendre un cycle cellulaire classique lorsqu'ils sont extraits des nodosités et étalés sur boite. Il est probable que les seules bactéries cultivables extraites des nodosités d'IRLC telles que *M. truncatula* sont celles qui ne sont pas différenciées et présentes dans les cordons d'infection.

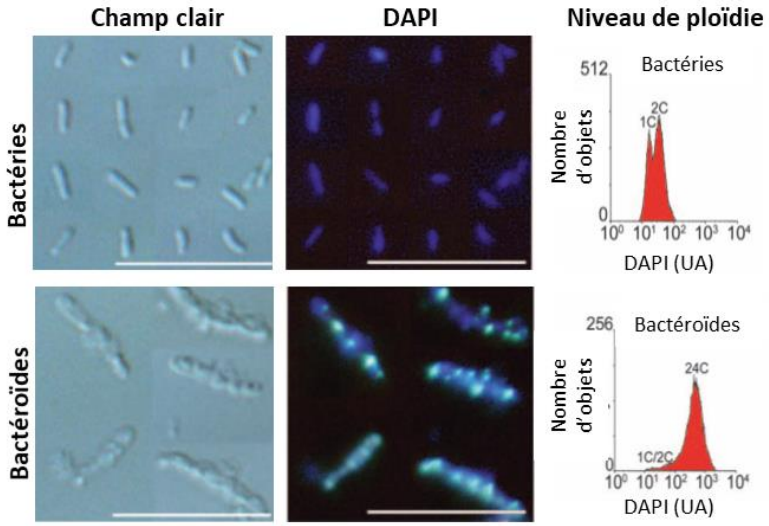


Figure 30: Comparaison de la taille et du niveau de ploïdie des cellules d'*Ensifer meliloti* 1021 en vie libre ou en symbiose avec *M. truncatula*. Les bactéries ont une ploïdie variant de 1 à 2C selon leur état dans le cycle cellulaire, c'est à dire avant ou après la phase S de réplication de l'ADN. Les bactéroïdes fixateurs d'azote matures et différenciés peuvent atteindre jusqu'à 24 fois leur contenu initial en ADN. Les bactéroïdes ont été extraits à 3 semaines post inoculation. La barre d'échelle indique 10µm. L'intensité de fluorescence de l'agent intercalant de l'ADN (DAPI) est mesurée en unités arbitraires (d'après Mergaert *et al.*, 2006).

3.3.2. Mécanismes moléculaires responsables de la différenciation : les peptides NCR

Les mécanismes responsables de ce processus de différenciation des bactéroïdes chez les IRLC ont été activement recherchés à partir des années 2000. Une famille de gènes a attiré une attention toute particulière, du fait de sa haute spécificité d'expression et de la diversité de ses représentants. Ces gènes codent des peptides riches en cystéines, présentant une structure proche de celle de peptides anti-microbiens. Ils sont appelés peptides NCR pour *Nodule-specific Cystein Rich*. Il en existe plus de 600 représentants chez *Medicago truncatula* (Mergaert *et al.*,

2003 ; Montiel *et al.*, 2017). Les peptides NCR possèdent un peptide signal peu conservé, des motifs typiques dotés de 4 ou 6 résidus cystéines, tandis que le reste de la séquence peptidique est hautement divergente, mais contribue à la charge électrique des peptides (Mergaert, 2018). Il existe deux types de motifs différents ressemblant aux défensines végétales ou à des neurotoxines animales (Mergaert *et al.*, 2003 ; Figure 30).

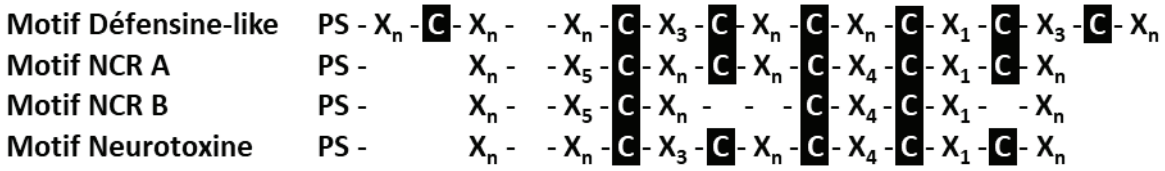


Figure 31: Similitudes dans la structure primaire entre les peptides NCR et des peptides antimicrobiens de type défensine. PS : Peptide signal. C : cystéine. X_n : n acides aminés autres que cystéines (d'après Mergaert *et al.*, 2003).

D'autres peptides NCR similaires à ceux du genre *Medicago* ont également été découverts chez 10 autres plantes appartenant à l'IRLC telles que *Pisum sativum*, *Galega orientalis*, *Astragalus canadensis* ou même *Glycyrrhiza uralensis* (Mergaert *et al.*, 2003 ; Montiel *et al.*, 2017). On peut d'ailleurs observer une corrélation positive entre le nombre de peptides NCR trouvés dans un génome de plante hôte et le niveau d'allongement de ses symbiontes (Montiel *et al.*, 2017). Il semblerait également qu'il existe une corrélation positive entre le niveau d'expression des NCR cationiques et le niveau d'allongement des bactéroïdes. Dans les nodosités de *M. truncatula* où les bactéroïdes subissent le plus fort allongement, l'expression des NCR cationiques représente 40% de l'expression totale des NCR (Roux *et al.*, 2014 ; Montiel *et al.*, 2017).

L'origine des peptides NCR remonte probablement à des duplications et réarrangements de peptides antimicrobiens ancestraux suivis par une diversification fonctionnelle (Mergaert, 2018). Ils pourraient également avoir évolué *de novo* à partir de régions génomiques non codantes utilisées comme matière première menant à l'apparition de nouveaux gènes (Mergaert, 2018). L'environnement génomique des NCR comprend à proximité immédiate des éléments transposables chez *M. truncatula.* Ces derniers ont un rôle possible dans la duplication des gènes qui est basée sur une recombinaison non homologue et une rétrotranscription des gènes à proximité du rétrotransposon (Mergaert, 2018). Cette duplication médiée par les

rétrotransposons est probablement à l'origine de l'explosion de la famille des NCR dans les génomes d'IRLC, et en particulier chez *M. truncatula*.

Il a été démontré que les peptides NCR sont dirigés vers la voie sécrétoire (Mergaert *et al.*, 2003). Par ailleurs, un crible génétique a été effectué chez *Medicago truncatula* et a permis l'identification de 7 mutants végétaux incapables de fixer l'azote, appelés mutants *dnf* (pour *defective in nitrogen fixation* ; Starker *et al.*, 2006). Le mutant *dnf1* est affecté dans la production de la protéine DNF1 exprimée dans les nodosités et faisant partie du complexe SPC (pour Signal Peptidase Complex) responsable de l'hydrolyse des peptides signaux dont ceux des peptides NCR. Elle est essentielle au processus de différenciation des bactéroïdes (Wang *et al.*, 2010). De façon concomitante, il a été prouvé que les peptides NCR sont responsables du processus de différenciation des bactéroïdes (Van de Velde *et al.*, 2010). Une colocalisation des peptides NCR avec les symbiosomes a été observée chez le génotype sauvage de *M. truncatula*, alors que ceux-ci restent au niveau du réticulum endoplasmique, et donc non sécrétés, chez le génotype *dnf1* (Van de Velde *et al.*, 2010, Figure 32). Cette hydrolyse des peptides NCR initie leur sécrétion dans les symbiosomes où ils provoquent le processus de différenciation des bactéroïdes.

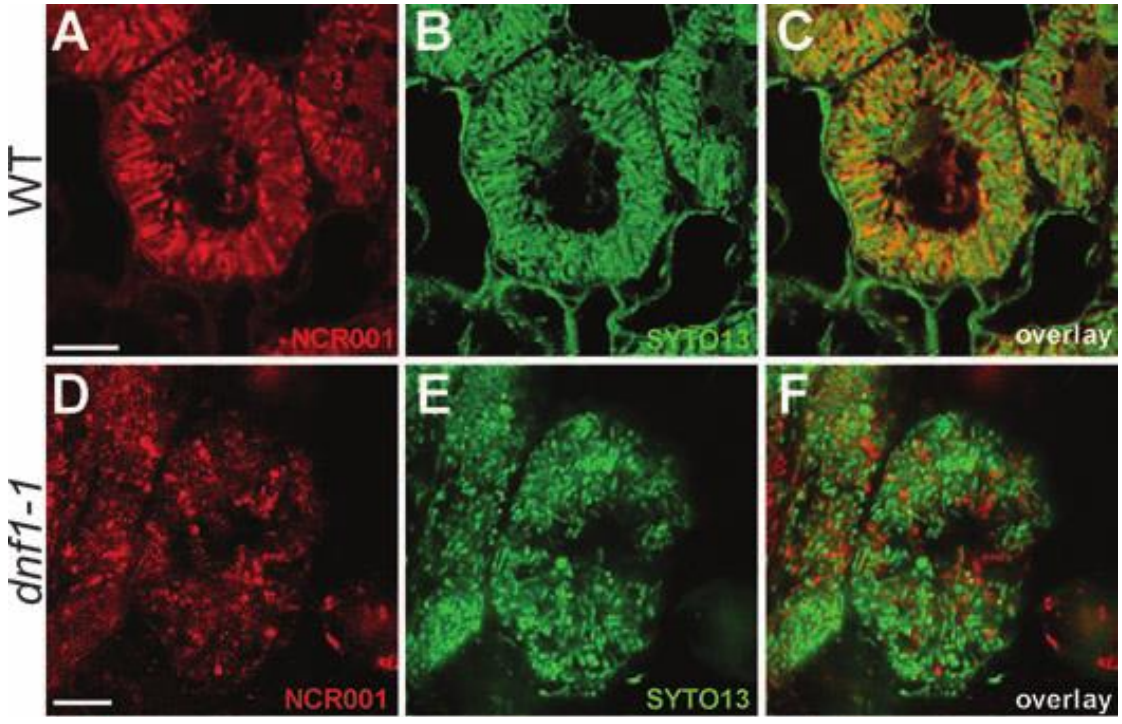


Figure 32: NCR001 colocalise avec les symbiosomes chez le génotype sauvage de *M. truncatula*, mais pas chez le mutant *dnf1*. (A, D) Immunolocalisation de NCR001. (B, E) SYTO13 (un agent intercalant fluorescent de l'ADN) marquant les bactéroïdes. (C, F) Superposition des images précédentes en condition WT (A, B, C) ou *dnf1* (D, E, F). Le mutant *dnf1* ne présente pas bactéroïdes différenciés (d'après Van de Velde *et al.*, 2010).

De plus, le traitement de *Ensifer meliloti* par des peptides NCR induit un début d'endoréplication des bactéries en vie libre (Van de Velde *et al.*, 2010 ; Figure 33). Le rôle des peptides NCR dans le processus de différenciation terminale précédant la fixation effective de l'azote atmosphérique a depuis été affiné. En effet, parmi les mutants *dnf* découverts autres que celui muté dans la signal-peptidase DNF1, deux d'entre eux impliquent des peptides NCR.

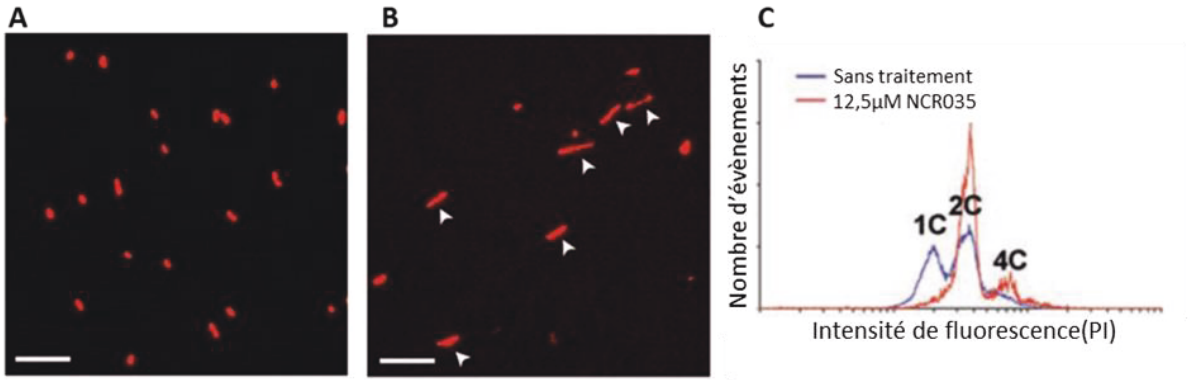


Figure 33: Les peptides NCR provoquent des caractéristiques de différenciation terminale en vie libre. (A, B) Cellules de *Ensifer meliloti* contrôle ou traitées avec des NCR035 qui induisent un allongement cellulaire. Les bactéries sont colorées avec de l'iodure de propidium. (C) Le niveau de ploïdie des bactéries traitées avec les peptides NCR035 est plus élevé que les bactéries non-traitées, mesuré par cytométrie en flux (d'après Van de Velde *et al.*, 2010).

Les mutants *dnf4* et *dnf7* ont subi la disruption des gènes codant respectivement les peptides NCR211 et NCR169 (Kim *et al.*, 2015 ; Horváth *et al.*, 2015). L'absence du NCR169 ou du NCR211 mène à une sénescence précoce des nodosités, et à la mort des bactéroïdes respectivement avant et après leur différenciation. Ces découvertes sont d'autant plus surprenantes du fait de la très grande diversité des peptides NCR. Leur nombre dépassant les 600 membres laissait supposer une possible une redondance fonctionnelle entre eux. Le phénotype drastique provoqué par l'absence d'un seul membre de la famille est donc assez contre-intuitif. Cela traduit l'immense complexité de la régulation des processus de différenciation et de mise en place de la symbiose fixatrice d'azote entre les IRLC telles que *M. truncatula* et ses endosymbiontes. Cette complexité se heurte également à des observations contradictoires concernant l'effet antimicrobien des peptides NCR *in vitro* et la nécessaire production d'au moins deux d'entre eux pour la survie des bactéroïdes et leur différenciation. Certains peptides NCR jouent également un rôle dans le choix du partenaire symbiotique. En effet, la gamme d'endosymbiontes fixant l'azote de manière efficace avec une IRLC donnée dépend de son répertoire de peptides NCR. C'est notamment le cas des peptides NCR NFS1 et NFS2. Les allèles NFS1$^{-/-}$ et NFS2$^{-/-}$ du génotype A17 bloquent la fixation effective d'azote par

les souches Rm41 (Wang *et al.*, 2017) et A145 (Yang *et al.*, 2017), ce qui n'est pas le cas pour les allèles de ces mêmes gènes chez le cultivar DZA315 donnant des nodosités Fix^+. L'effet d'autres peptides NCR fait l'objet d'intenses recherches actuellement. C'est le cas du peptide NCR247, qui a été particulièrement bien caractérisé au niveau cellulaire.

3.3.3. Les protéines bactériennes impliquées dans la différenciation des bactéroïdes

Une première étude a montré l'internalisation du NCR247 par la bactérie *in vitro* et identifié ses cibles potentielles, ou du moins les protéines interagissant physiquement avec le peptide, par chromatographie d'affinité (Farkas *et al.*, 2014). Deux cibles intéressantes ont été identifiées comme interagissant avec le NCR247 : la protéine FtsZ dont la polymérisation est impliquée dans la formation de l'anneau-Z requis pour la division cellulaire ; ainsi que la chaperonne moléculaire GroEL1. L'interaction FtsZ-NCR247 pourrait inhiber la polymérisation de FtsZ empêchant la division cellulaire, prenant ainsi potentiellement part au processus de différenciation lié à un allongement forcé. De plus, GroEL1 est requise pour une symbiose fonctionnelle entre *E. meliloti* et *M. truncatula* (Farkas *et al.*, 2014).

Des approches transcriptomiques (Penterman *et al.*, 2014) ont montré l'effet majeur du peptide NCR247 *in vitro* sur les gènes impliqués dans la régulation du cycle cellulaire bactérien. Par exemple, les niveaux d'expression des régulateurs « maître » *dnaA* et *ctrA* chutent drastiquement lors d'un traitement avec ce peptide. Chez les alphaprotéobactéries, DnaA coordonne l'initiation de la réplication de l'ADN, mais agit également comme un facteur de transcription activateur de la division cellulaire (Hottes *et al.*, 2005). CtrA est un régulateur de réponse actif lorsque phosphorylé, induisant la division cellulaire et inhibant la réplication de l'ADN entre autres chez *Ensifer meliloti* et *Caulobacter crescentus* (Biondi *et al.*, 2006 ; Pini *et al.*, 2015). Leur faible expression inhibe ainsi la division cellulaire et induit la réplication de l'ADN. Une expression décroissante de CtrA est d'ailleurs observée avec la progression de la différenciation des bactéroïdes dans les nodosités de *M. truncatula* (Roux *et al.*, 2014). Bien que la mutation du gène soit létale, des souches d'*Ensifer meliloti* exprimant faiblement le gène *ctrA* ont été obtenues et présentent des phénotypes variés. Les faibles niveaux de CtrA active mènent à une morphologie similaire aux bactéroïdes différenciés *in vitro*. De plus, un mutant *ΔdivJ* (un régulateur négatif de CtrA) dont l'effet supposé est une accumulation importante de

protéine CtrA active, montre un développement des nodosités perturbé (Pini et al, 2013 ; Pini *et al.*, 2015, Figure 34). L'interaction entre des peptides NCR et CtrA n'a pas été démontrée et n'est probablement pas directe étant donnée le nombre de protéines impliqués dans la régulation de son activité.

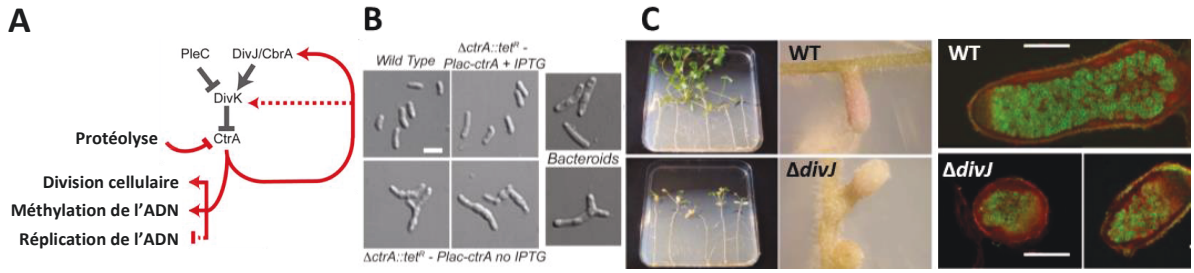


Figure 34: La différenciation des bactéroïdes fait intervenir le régulateur « maitre » CtrA. (A) Fonctionnement de la boucle de régulation de CtrA et ses effets sur le cycle cellulaire. (B) Les faibles doses de CtrA cellulaire mènent à une morphologie similaire aux bactéroïdes *in vitro*. La souche comparée à la souche sauvage est un mutant Δ*ctrA* dont l'expression de *ctrA* est inductible en présence d'IPTG. (C) Phénotype d'un mutant Δ*divJ*, dont l'effet attendu est une accumulation de protéine CtrA active, montrant un développement des nodosités perturbées (d'après Pini *et al.*, 2013 ; Pini *et al.*, 2015).

Un autre gène indispensable à la différenciation des bactéroïdes a été identifié bien avant même la découverte des peptides NCR chez *E. meliloti*. Il a été nommé *bacA* pour *bacteroid development A*. *bacA* code une protéine membranaire ressemblant aux perméases des transporteurs ABC (pour « ATP-Binding Cassette ») (Glazebrook *et al.*, 1993). Les transporteurs ABC fonctionnent habituellement sous forme de trois sous-unités : une sous-unité PBP (*periplasmic binding protein*) qui se lie à son ligand spécifique et peut ensuite interagir avec deux sous-unités transmembranaires formant une perméase. Cette perméase est un canal membranaire dont les changements de conformation provoquent le transport actif du ligand assurés par la sous-unité ATPase cytosolique (Davidson *et al.*, 2008). Le gène *bacA* ne possède plus que la sous-unité perméase. Les deux autres sous-unités, si elles existent, ne sont pas présentes à proximité directe de l'environnement génomique de *bacA* comme c'est souvent le cas chez les transporteurs ABC. La protéine BacA présente 64% d'identité avec un transporteur de peptides antimicrobiens de *Escherichia coli* appelé SbmA (Glazebrook *et al.*, 1993). Les bactéroïdes Δ*bacA* présentent un phénotype d'absence de différenciation et de mort cellulaire rapide après leur relargage dans le cytoplasme, et les nodosités subissent une sénescence

prématurée caractérisée par une absence de leghémoglobine (Glazebrook *et al.*, 1993 ; Haag *et al.*, 2011, Figure 35).

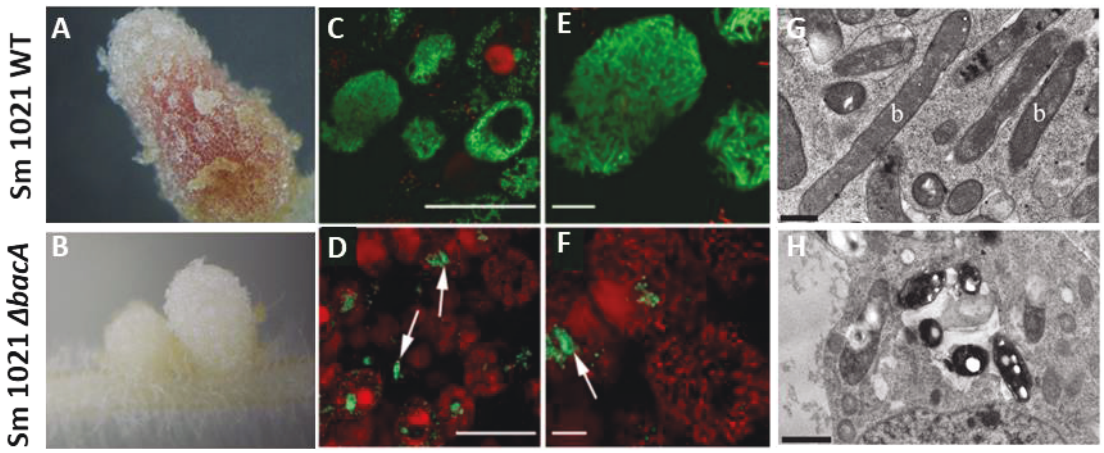


Figure 35: Phénotype symbiotique des nodosités infectées par le mutant bactérien *ΔbacA* d'*E. meliloti*. Les génotypes WT (A, C, E, G) et *ΔbacA* (B, D, F, G) sont comparés. (A, B) Morphologie des nodosités, montrant l'avortement de la nodosité et l'absence de leghémoglobine chez le mutant *ΔbacA* (B). (C, D, E, F) Images de microscopie confocale avec marquage live/dead (syto9 et iodure de propidioum) marquant respectivement les bactéroïdes vivants en vert et morts en rouge. Les bactéroïdes du mutant *ΔbacA* ne présentent pas d'allongement et sont morts, en dehors de ceux présents dans les cordons d'infections pointés par les flèches blanches (D, F). (E, F) Agrandissements respectifs de (C, D). (G, H) Comparaison au microscope électronique à transmission des bactéroïdes sauvages et *ΔbacA*, montrant l'absence d'allongement des bactéroïdes et la présence de grains de polyhydroxybutyrate, peu denses aux électrons (d'après Arnold *et al.*, 2013 ; Haag *et al.*, 2011).

De façon remarquable, des homologues de *bacA* ont été découverts chez de nombreuses bactéries, dont des pathogènes responsables d'infections intracellulaires d'Eucaryotes telles que *Brucella abortus* (LeVier *et al.*, 2000) et *Mycobacterium tuberculosis* (Domenech *et al.*, 2009) où ils sont nécessaires à l'établissement d'une infection chronique chez leur hôte (LeVier *et al.*, 2000 ; Domenech *et al.*, 2009 ; Haag *et al.*, 2013). De plus, il est possible de complémenter au moins partiellement la délétion du gène *bacA* d'*E. meliloti* par ceux de *B. abortus* (Wehmeier *et al.*, 2010) et *M. tuberculosis* (Arnold *et al.*, 2013). Cette complémentation conduit à la formation de nodosités roses différentes de celles élicitées par le mutant *ΔbacA,* mais pas aussi riches en leghémoglobine que la souche sauvage. En conditions symbiotiques, le rôle de BacA est relié à celui des peptides NCR. En effet, lorsqu'on inocule le mutant plante *DNF1* avec la souche mutante *ΔbacA*, on n'observe ni différenciation (car pas d'adressage des NCR aux symbiosomes) mais également une absence de mort cellulaire des bactéroïdes (Haag *et al.*, 2011). *In vitro*, BacA participe à la protection contre les NCR cationiques (Haag *et al.*, 2011). De plus, BacA est nécessaire chez les IRLC productrices de NCR mais pas chez les légumineuses de la famille des phaséoloïdes qui n'en produisent pas (Karunakaran *et al.*, 2010,

Maruya et Saeki, 2010). L'état actuel des connaissances concernant le processus de différenciation des bactéroïdes est résumé par la Figure 36 (Alunni et Gourion, 2016).

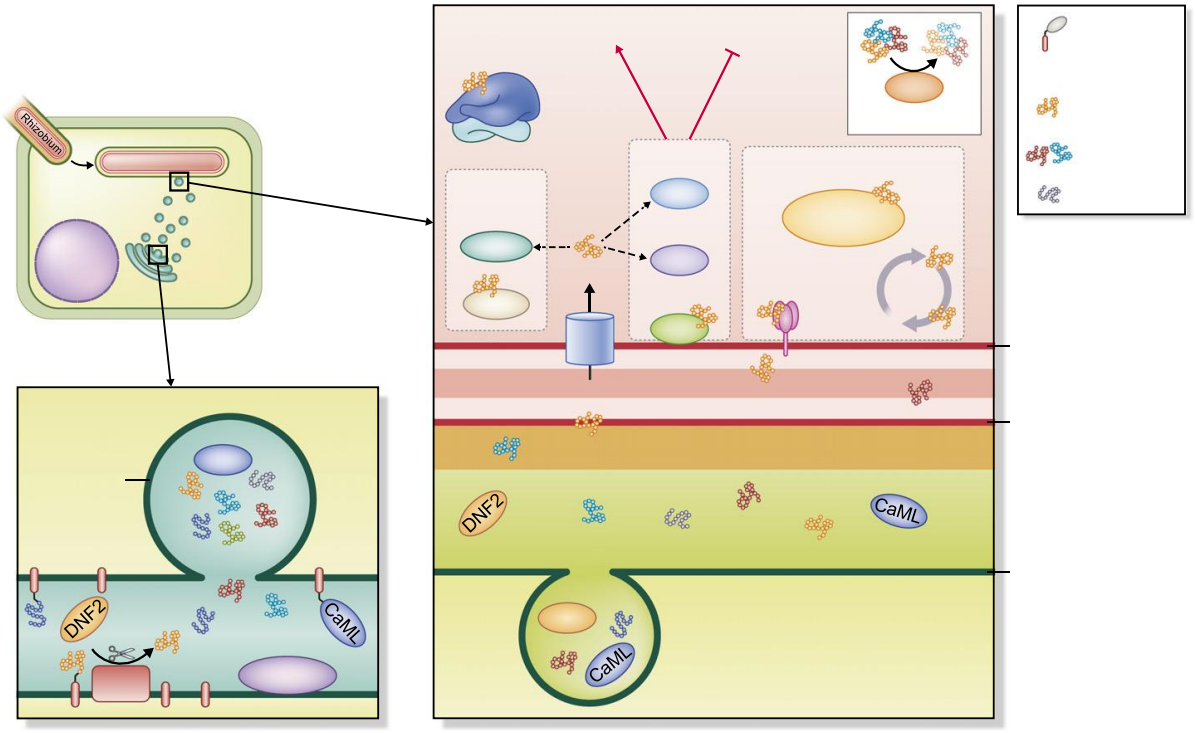


Figure 36: Mécanismes moléculaires de la mise en place lors du processus de différenciation des bactéroïdes chez *E. meliloti*. La sécrétion des peptides NCR de l'appareil de Golgi aux bactéroïdes la signal-peptidase DNF1, et nécessite les protéines VAMP721s pour permettre la fusion vésiculaire au niveau de la membrane des symbiosomes. LG : Lumière de l'appareil de Golgi, CB : cytosol bactérien, PG : peptidoglycane, LPS : lipopolysaccharide, EPB : espace péri-bactéroïdien, CP : cytosol de la cellule symbiotique de plante (d'après Alunni et Gourion, 2016).

Le processus de différenciation des bactéroïdes est crucial, mais n'est pas spécifique aux IRLC. Ce travail de thèse s'intéresse plus spécifiquement à une autre association symbiotique dont les caractéristiques/ spécificité ont été découvertes plus récemment que les travaux initiaux sur l'association *Ensifer-Medicago*. Ce travail porte plus particulièrement sur les bactéries du genre *Bradyrhizobium*, qui forment une symbiose avec des légumineuses des familles des *Phaseolides* (e.g *Glycine max*) et des Dalbergioïdes (e.g *Aeschynomene* spp.) (Figure 37).

4. La symbiose *Bradyrhizobium-Aeschynomene*

4.1. Diversité des *Aeschynomene* et établissement de groupes d'inoculation croisés

Le genre *Aeschynomene* appartient au clade des Dalbergioïdes qui a émergé il y a 55 Ma (Chaintreuil *et al.*, 2016). Il contient actuellement environ 180 espèces. Ces espèces sont, pour la moitié d'entre elles natives du continent américain et l'autre moitié des régions subtropicales, principalement d'Asie et d'Afrique. Ce genre comprend des plantes pérennes et annuelles telles que des herbacées et des arbustes dont de nombreuses espèces sont présentes dans les savanes et les forêts sèches. Certaines espèces d'*Aeschynomene* correspondent, en revanche, à des plantes semi-aquatiques. Elles se développent dans des marais, le long des cours d'eau ou des rizières. Elles partagent la particularité d'arborer également des nodosités caulinaires avec d'autres légumineuses semi-aquatiques du genre *Sesbania* et *Neptunia* (Alazard, 1985). Les *Aeschynomene* développent un type de nodosité appelé Aeschynomenoïde. C'est une synapomorphie des Dalbergioïdes qui se distingue des autres nodosités déterminées car la quasi-intégralité des cellules de la nodosité sont infectées et elles sont également toujours associées à une racine latérale ou des racines adventives pour les nodosités racinaires et caulinaires respectivement (Lavin *et al.*, 2001, Figure 37). Les nodosités d'*Aeschynomene* sont colonisées par des bactéries du genre *Bradyrhizobium*, dont différents groupes d'inoculation croisés ont été déterminés (Alazard, 1985). Un groupe d'inoculation croisé se définit comme un ensemble d'hôtes capables d'établir une symbiose fixatrice d'azote avec le même ensemble de symbiontes, et réciproquement. Cette analyse initiale s'est basée sur l'inoculation de 26 espèces d'*Aeschynomene* par 15 souches de *Bradyrhizobium* isolées de nodosités racinaires et caulinaires d'*Aeschynomene* en conditions naturelles (Alazard, 1985). Trois groupes d'inoculation croisés végétaux et cinq groupes d'inoculation bactériens ont émergé de ces travaux, avec des souches « modèles » représentatives de chaque groupe (Alazard, 1985 ; Chaintreuil *et al.*, 2013 ; Figure 38).

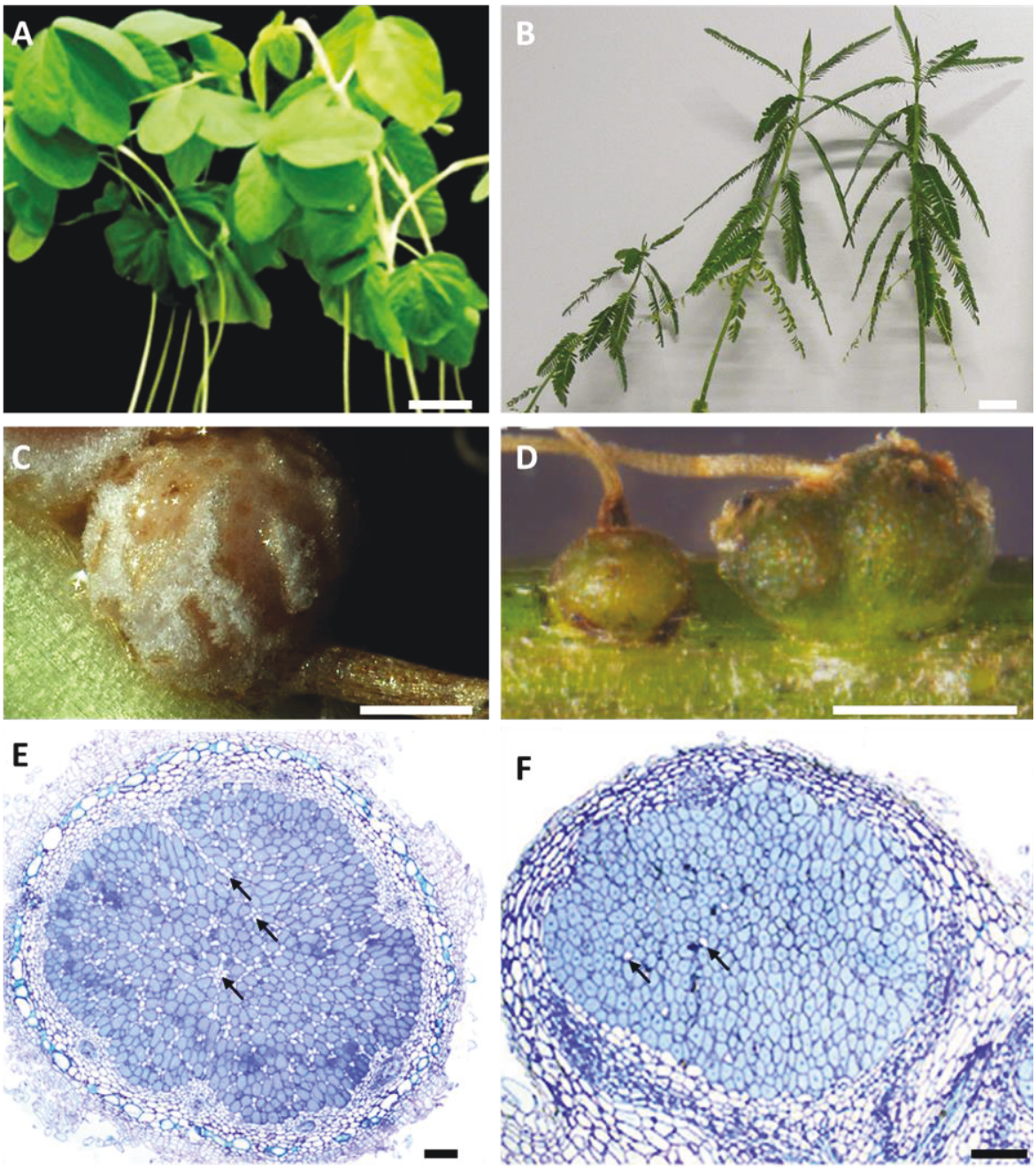


Figure 37: Comparaisons phénotypiques et histologiques entre les nodosités de soja (*Glycine max*, gauche) et *Aeschynomene indica* (droite). (A, B) Photographies de parties aériennes. Barres d'échelle : 3,5 cm. (C, D) Observation au macroscope de nodosités de soja infectées par *B. diazoefficiens* USDA110 (C) et de *A. indica* infectées par *Bradyrhizobium* sp. ORS285 (D). Les nodosités sont dans les deux cas déterminées. Barres d'échelle : 1 mm. (E, F) Coupes de nodosités colorées au bleu de toluidine, montrant très peu de cellules non-infectées (blanches et indiquées par des flèches noires) dans la nodosité Aeschymenoïde de *A. indica* comparé à celle de soja. Barres d'échelle : 100µm (d'après Barrière *et al.*, 2017 (A, E) ; photographies personnelles (B, C), Busset *et al.*, 2016 (D) et Guefrachi *et al.*, 2015 (F)).

Le groupe I correspond à des espèces d'*Aeschynomene* à nodulation uniquement racinaire en association avec des bradyrhizobia non-photosynthétiques à spectre d'hôte assez large comme par exemple *B. diazoefficiens* USDA110 qui infecte aussi les nodosités de *Phaseolides*

comme le soja, *Macroptilium atropurpureum* et *Vigna unguiculata* (Koch *et al.*, 2010). Le groupe III est représenté par des plantes à nodulation caulinaire et racinaire, comme *A. evenia* ou *A. indica*. Elles peuvent être nodulées par des bradyrhizobia photosynthétiques (ORS278 et BTAi1) ou non (STM3844) uniquement capables d'induire des nodosités avec ce groupe. Le groupe II dont fait partie *A. afraspera*, peut être nodulé à la fois par des *Bradyrhizobium* non-photosynthétiques nodulant également le groupe I (e.g USDA110) ou non (e.g ORS354), et des *Bradyrhizobium* photosynthétiques qui nodulent également le groupe III (e.g ORS285).

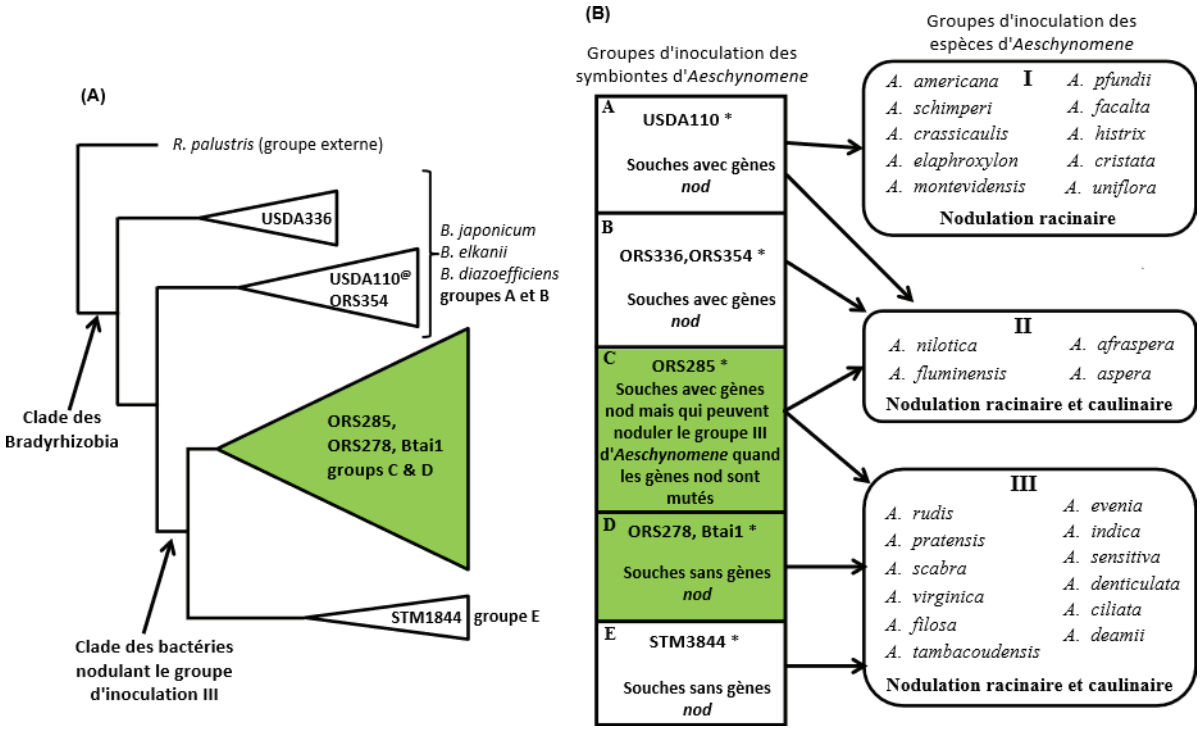


Figure 38: Groupes d'inoculation croisés entre bradyrhizobia symbiotiques et leurs hôtes *Aeschynomene* compatibles. (A) Relations phylogénétiques entre les souches. (B) Groupes d'inoculation bactériens définis parmi les diverses espèces d'*Aeschynomene*. Les encadrés à fond vert désignent les espèces bactériennes photosynthétiques. * : espèce représentative du groupe (d'après Chaintreuil *et al.*, 2013).

4.2. Les processus de nodulation indépendante des facteurs Nod chez les *Aeschynomene*

4.2.1. Découverte d'un processus symbiotique Nod-indépendant

Une particularité des bactéries du genre *Bradyrhizobium* est que certaines souches sont capables de réaliser des symbioses fixatrices d'azote sans posséder les gènes *nod*. Les facteurs

Nod ont longtemps été considérés comme essentiels à l'établissement de la symbiose rhizobium légumineuses, car l'ensemble des rhizobia testées jusqu'en 2007 possédaient des facteurs Nod.

De plus, il est plus complexe de démontrer l'absence de ces facteurs que leur présence par spectrométrie de masse. L'avènement de la génomique appliquée à de nouveaux modèles de rhizobia a permis de lever ce verrou technologique en faisant des séquençages massifs. C'est notamment en obtenant le génome de deux souches de *Bradyrhizobium* photosynthétiques, ORS278 et BTAi1, que l'absence des gènes *nod* a pu être observée (Giraud *et al.*, 2007). Les *Aeschynomene* du groupe III peuvent donc établir des nodosités fixatrices d'azote en l'absence de facteurs Nod, par un mécanisme encore inconnu aujourd'hui. L'existence de souches du groupe d'inoculation C, possédant également des facteurs Nod et capables de noduler les plantes du groupe III, a aussi permis de démontrer l'existence de symbioses Nod-indépendantes. En effet, une souche *ΔnodB* de *Bradyrhizobium* sp. ORS285 perdant la capacité de synthétiser des facteurs Nod perd sa capacité à noduler les *Aeschynomene* du groupe II, mais pas celles du groupe III (Giraud *et al.*, 2007 ; Figure 39).

Figure 39: Phénotypes de nodulation des souches ORS278, ORS285 et ORS285*ΔnodB* avec *A. afraspera* et *A. sensitiva*. (A) *A. afraspera* (groupe II) et (B) *A. sensitiva* (groupe III). NI : condition contrôle non inoculée (d'après Giraud *et al.*, 2007).

4.2.2. Les processus Nod-indépendants et T3SS-dépendant

Quelques années après la découverte du processus précédemment décrit, un processus alternatif d'initiation de l'infection a été découvert entre des génotypes de soja et de *Bradyrhizobium elkanii* USDA61. L'initiation de l'interaction entre le soja et cette souche est

Nod-dépendante. Cependant, des nodosités ont tout de même été observées chez un génotype particulier de soja (En1282) muté dans le gène *nfr1* de perception des facteurs Nod. Cette complémentation de la signalisation des facteurs Nod n'est plus observée lorsque la souche USDA61 est mutée pour son système de sécrétion de type III (T3SS). Il semblerait donc qu'une voie alternative faisant intervenir des effecteurs de type III puisse également initier l'interaction symbiotique (Okazaki *et al.*, 2013 ; Figure 40A). Des études transcriptomiques ont montré que des effecteurs potentiels de *Bradyrhizobium elkanii* USDA61 activent les gènes ENOD40 et NIN, qui sont des gènes symbiotiques en aval de la voie de signalisation des facteurs Nod (Okazaki *et al.*, 2013 ; Oldroyd, 2013). Ce résultat était surprenant car la plupart des bradyrhizobia photosynthétiques comme la souche ORS278 ne produisent pas de facteurs Nod et ne possèdent pas non plus de T3SS (Okazaki *et al.*, 2015). Plus tard, il a été montré que des souches de *B. elkanii* sont capables de noduler *Aeschynomene indica* du groupe III. Les phénotypes de nodulation de *A. indica* ont donc été comparés avec les souches sauvage, ΔT3SS et *Δnod* de *B. elkanii* USDA61. Les souches sauvage et *Δnod* sont capables d'initier la formation de nodosités, mais pas la souche ΔT3SS (Figure 40B). Ces résultats montrent qu'il existerait au moins deux processus de nodulation indépendants des facteurs Nod, respectivement dépendant et indépendant de l'utilisation du T3SS (Okazaki *et al.*, 2015).

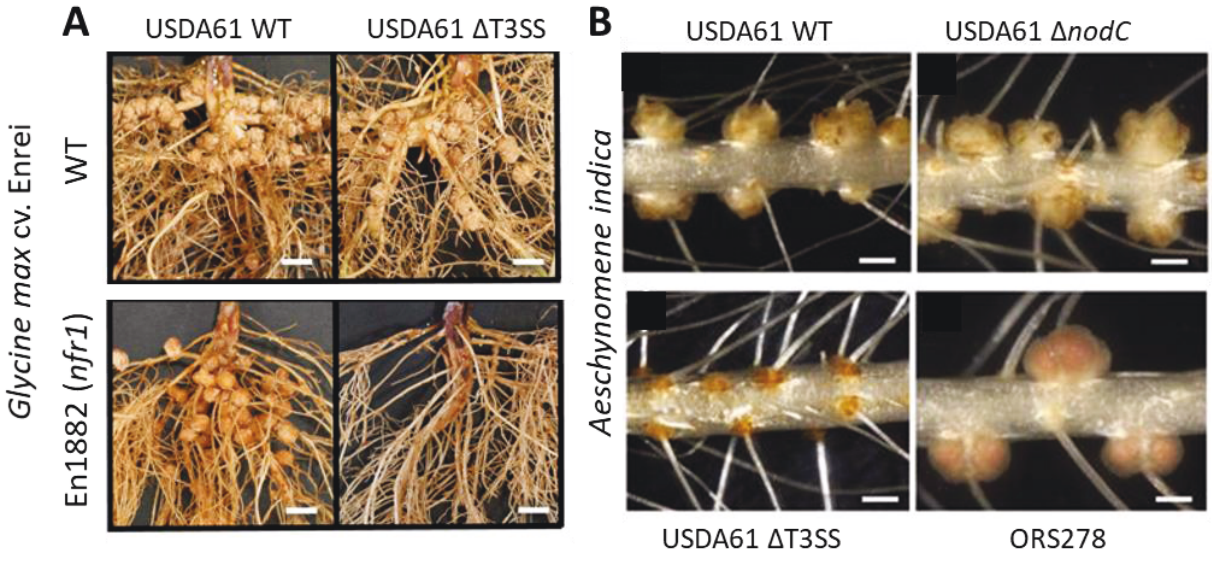


Figure 40: Rôle du système de sécrétion de type III dans l'initiation de l'interaction symbiotique entre *Bradyrhizobium elkanii* USDA61 et ses plantes hôtes. (A) Phénotypes de nodulation de deux génotypes de *Glycine max* cv. Enrei avec les souches WT et ΔT3SS à 30 jours post-inoculation (dpi). L'association *nfr1* et ΔT3SS ne donne pas de nodosités. Barre d'échelle : 4mm (B) Phénotypes de nodulation de *A. indica* avec les souches WT, Δ*nodC* et ΔT3SS de *B. elkanii* et la souche de *Bradyrhizobium* sp. ORS278. L'absence de nodosités n'est observée que lors de l'inoculation par le mutant ΔT3SS. Barres d'échelle : 1mm (d'après Okazaki *et al.*, 2013 ; Okazaki *et al.*, 2015).

4.2.3. La signalisation Nod-indépendante active certains acteurs de la voie de signalisation commune des endosymbioses

Les molécules responsables du dialogue moléculaire lors de la symbiose Nod-indépendante T3SS-indépendante sont encore inconnues à ce jour, mais il est fort probable que cette symbiose utilise une partie de la voie symbiotique commune aux autres endosymbioses. Des orthologues des gènes symbiotiques clés ont été identifiés chez *Aeschynomene evenia* et leurs rôles dans la symbiose Nod-indépendante ont été testés (Fabre *et al.*, 2015). Des expériences d'ARN-interférence ont été effectuées sur 3 gènes : SymRK (DMI2) qui intervient très précocement dans la cascade de signalisation, CCaMK qui joue un rôle essentiel dans la coordination de l'infection et de l'organogénèse, et HK1 qui intervient en aval de CCaMK lors de la transduction du signal des facteurs Nod (Fabre *et al.*, 2015). Dans les trois cas, la diminution du nombre de nodosités observées est significative, indiquant le recrutement de ces trois gènes dans le processus symbiotique Nod-indépendant. L'analyse du transcriptome de nodosités d'*A. evenia* a permis de mettre en évidence un plus grand nombre de gènes de la voie symbiotique commune au cours de la symbiose Nod-indépendante (Chaintreuil *et al.*, 2016 ; Figure 41).

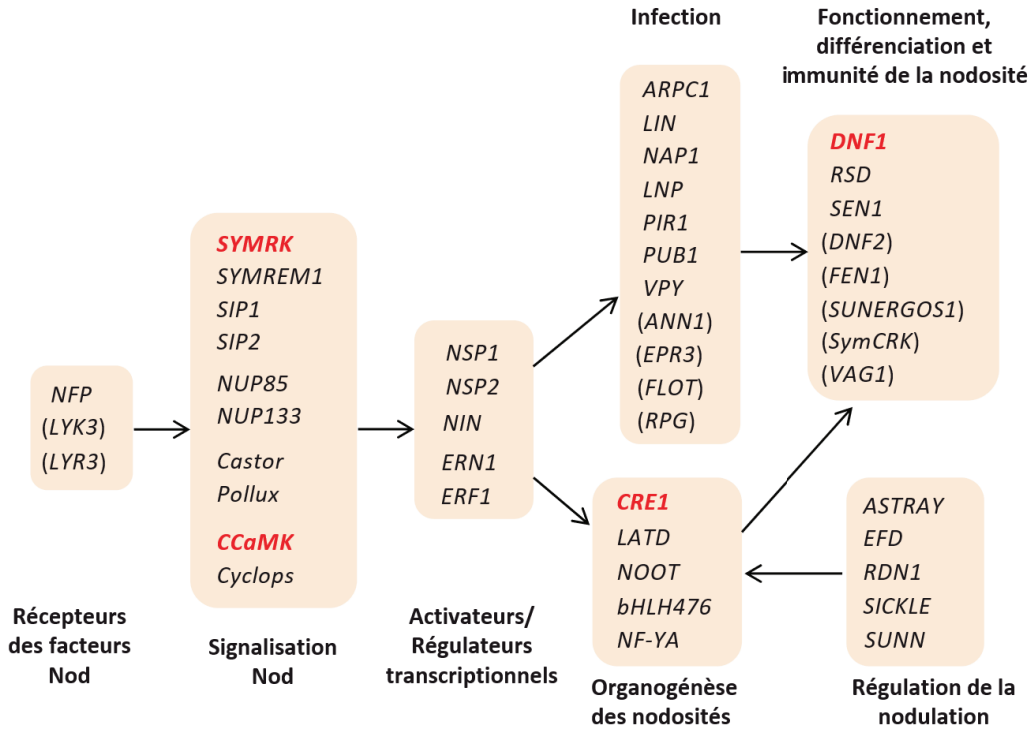


Figure 41: Modèle simplifié de la voie de signalisation symbiotique entre rhizobia et légumineuses. Les gènes symbiotiques identifiés chez *M. truncatula* ou *L. japonicus* sont listés et classés dans des fonctions symbiotiques. Les orthologues putatifs ont été recherchés dans le transcriptome d'*Aeschynomene evenia* (groupe III). Les gènes non identifiés sont entre parenthèse et ceux en rouge sont validés fonctionnellement comme requis pour la symbiose Nod-indépendante chez *A. evenia* (d'après Chaintreuil *et al.*, 2016).

De façon attendue, les protéines LYK3 et LYR3 à forte affinité pour les facteurs Nod ne sont pas retrouvées dans le transcriptome, pas plus que les gènes impliqués dans la formation du cordon d'infection et les étapes précoces de reconnaissance des exopolysaccharides bactériens comme le gène EPR3 (Chaintreuil *et al.*, 2016). En effet, le mode d'infection des *Aeschynomene* par les symbiontes diffère singulièrement de l'infection du poil absorbant caractérisé chez les plantes modèles de la famille des légumineuses.

4.3. Infection et organogénèse des nodosités chez les *Aeschynomene*

On se limitera ici à utiliser les exemples d'*Aeschynomene afraspera* et *Aeschynomene indica*, appartenant aux groupes d'inoculation II et III et utilisant un processus symbiotique Nod-dépendant et Nod-indépendant, respectivement. Elles peuvent être nodulées par une même espèce bactérienne, *Bradyrhizobium* sp. ORS285 appartenant au groupe d'inoculation bactérien C (Figure 38). L'infection se fait dans les deux cas selon le mode du crack-entry.

Concernant *A. indica*, l'infection et l'organogénèse des nodosités commencent par la colonisation massive de la surface et des espaces intercellulaires de poils axillaires situés à l'aisselle des racines secondaires (Bonaldi *et al.*, 2011 ; Figure 42A). Il n'y a pas contre aucune courbure du poil absorbant ni de formation de cordon d'infection comme dans les autres couples symbiotiques modèles (Bonaldi *et al.*, 2011). Les bactéries infectent et se divisent ensuite dans une première couche de cellules corticales qui meurent à la suite de l'infection de façon similaire à l'entrée d'une bactérie pathogène. Ces cellules mortes sont ensuite collapsées par la pression de turgescence des cellules voisines, déversant ainsi leur contenu sous forme de diverticules qui s'enfoncent plus profondément dans le cortex racinaire. Les bactéries infectent alors les cellules corticales plus profondes au contact des cellules collapsées. Ce deuxième type de cellules infectées va devenir ce qu'on appelle des cellules fondatrices capables de division et conservant leur intégrité. Elles mettront rapidement en place la nodosité de type Aeschynomenoïde qui est fonctionnelle dès 5 jours après inoculation (Figure 41B, C, F). Les bactéroïdes passent d'un morphotype indifférencié à un morphotype sphérique entre le 4ème et le 5ème jour post-inoculation (Figure 42D, E, F).

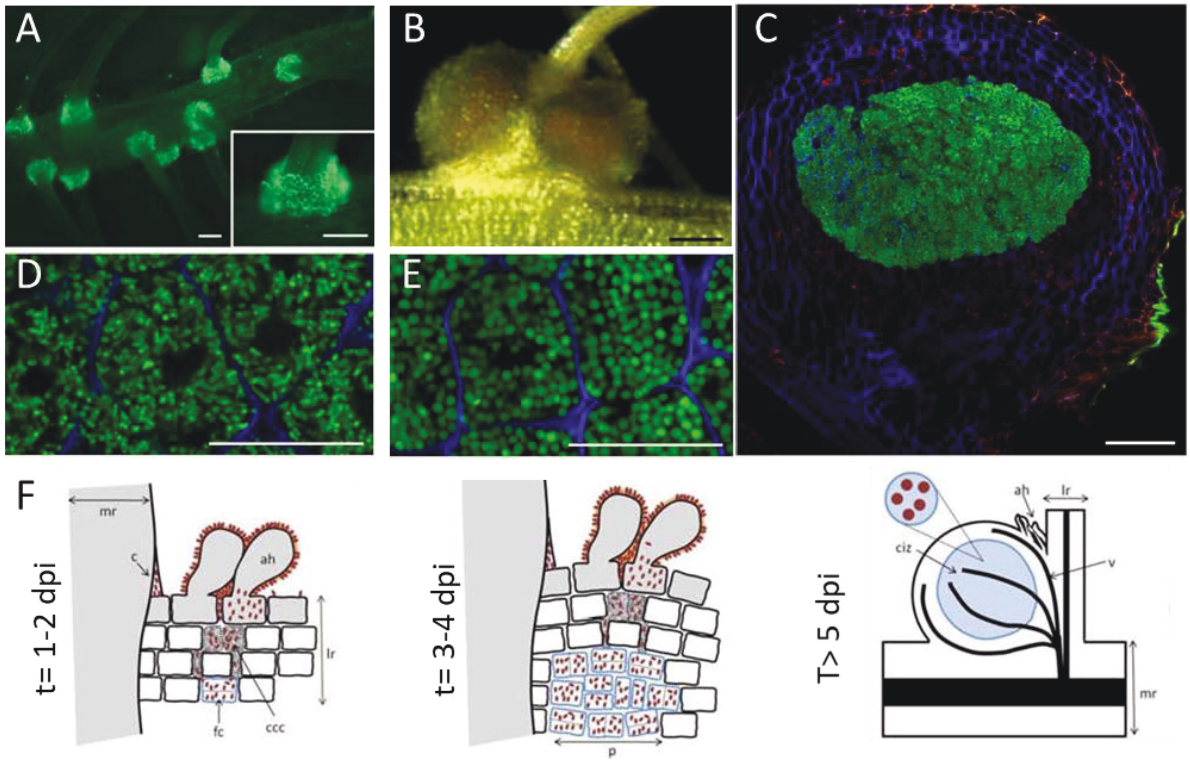


Figure 42: Infection et organogénèse des nodosités d'*Aeschynomene indica* inoculées par la souche *Bradyrhizobium* sp. ORS285. (A) Image stéréomicroscopique montrant la colonisation massive des poils axillaires des racines secondaires par la souche ORS285 *ΔnodB* transformée par un gène codant une GFP avec un agrandissement sur les poils axillaires d'une racine secondaire. Barre d'échelle : 250µm. (B) Nodosité mature à 10 dpi observée au macroscope. Barre d'échelle : 500µm. (C) Coupe de jeune nodosité à 6 dpi observée en microscopie confocale. Les cellules végétales de la nodosité sont quasiment toutes infectées. Barre d'échelle : 100µm. (D, E) Coupes de nodosités montrant la transition de morphotype des bactéroïdes passant de non différencié à sphérique entre 4 dpi (D) et 5 dpi (E). Barre d'échelle : 20µm. (F) Représentation schématique du processus d'organogénèse de la nodosité d'*Aeschynomene indica* (groupe III) qui effectue une symbiose Nod-indépendante. Abréviations : mr, racine principale ; lr, racine latérale ; c, crack ; ah, poil axillaire ; ccc, cellules corticales collapses ; p, primordium ; dc, cellule fondatrice du primordium de la nodosité ; v, vascularisation ; ciz, zone centrale infectée. Les cellules de l'épiderme sont colorées en gris et les cellules corticales en blanc (modifié d'après Bonaldi *et al.*, 2011).

Quelques différences peuvent être observées concernant le mode d'infection d'*A. afraspera* par rapport à celui de *A. indica*. Les poils axillaires sont beaucoup plus allongés et beaucoup moins nombreux que chez *A. indica* et ne sont pas autant colonisés par les bactéries (Bonaldi *et al.*, 2011 ; Figure 43A). On observe une pénétration intercellulaire des bradyrhizobia entre un poil axillaire de la racine secondaire et une cellule de l'épiderme de la racine primaire. L'émergence des racines secondaires provoque des zones de fissure de l'épiderme par lequel s'initie la pénétration des bactéries (Bonaldi *et al.*, 2011 ; Figure 43F). Les premières couches de cellules corticales sont infectées mais ne subissent pas une mort cellulaire comme chez *A. indica*. Au contraire, les bactéries s'y développent au sein de tubulures ressemblant à des cordons d'infection, et les cellules végétales deviennent géantes et formeront à terme une

excroissance au-dessus de la nodosité colonisée par des bactéries non-différenciées (Figure 43B, C, D). Les bactéries traversent aussi plus profondément (3 à 4 couches) le cortex racinaire de manière intercellulaire avant d'en infecter des cellules qui deviendront les cellules fondatrices de la nodosité (Bonaldi *et al.*, 2011, Figure 43F). Les bactéroïdes subissent un allongement par rapport aux bactéries présentes dans les cellules géantes de l'excroissance (Figure 43E).

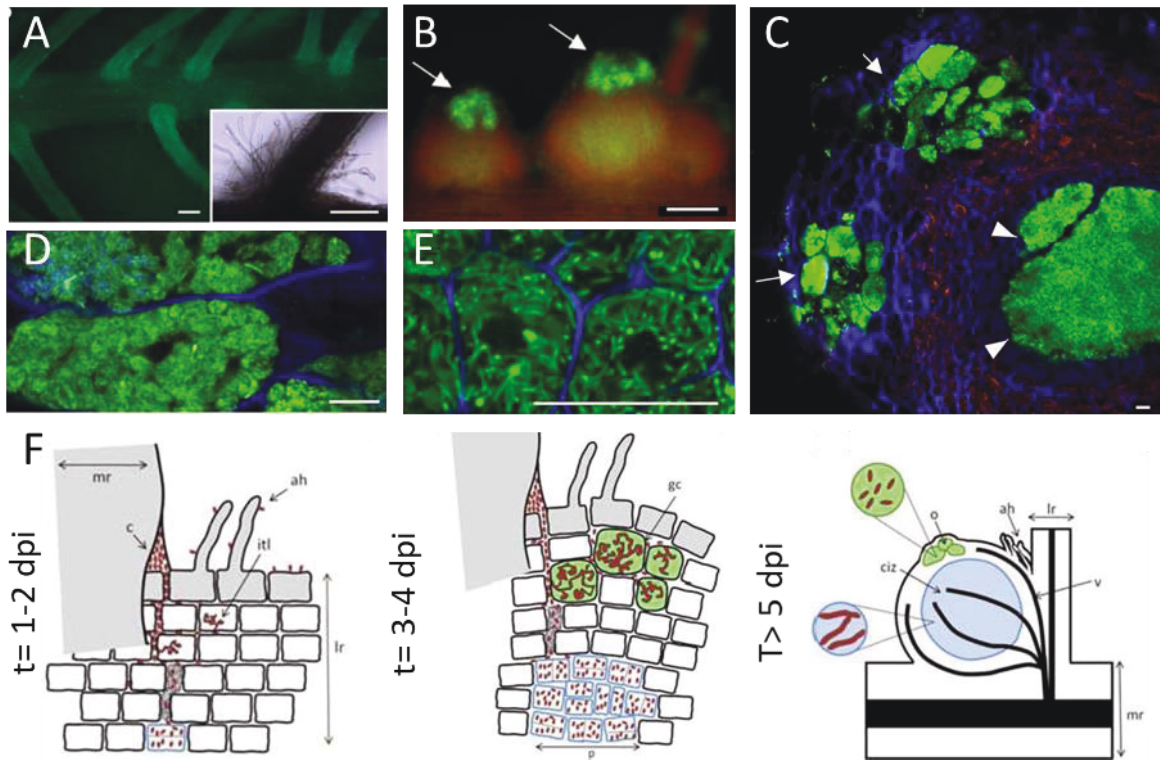


Figure 43: Infection et organogénèse des nodosités d'*Aeschynomene afraspera* inoculées par la souche *Bradyrhizobium* sp. ORS285. (A) Image stéréomicroscopique montrant la colonisation des poils axillaires des racines secondaires par la souche ORS285 transformée par un gène codant une GFP avec un insert zoomant sur les poils axillaires d'une racine secondaire. La colonisation est plus faible que celle observée chez *A. indica* Barres d'échelle : 250µm. (B) Image stéréomicroscopique d'une nodosité mature à 10 dpi où l'on observe les bactéries GFP^+ dans les excroissances. Barre d'échelle : 500µm. (C) Coupe de nodosité en microscopie confocale d'une jeune nodosité à 6 dpi. Les cellules végétales de la nodosité sont quasiment toutes infectées. Barre d'échelle : 20µm. (D,E) Coupes de nodosités montrant la différence de morphotype entre les bactéries des cellules géantes de l'excroissance (D) et les bactéroïdes allongés au sein des cellules symbiotiques (E). Barres d'échelle : 20µm. (F) Représentation schématique du processus d'organogénèse de la nodosité d'*Aeschynomene afraspera* (groupe II) qui effectue une symbiose dépendante des facteurs Nod. Abréviations : mr, racine principale ; lr, racine latérale ; c, crack ; ah, poil axillaire ; ccc, cellules corticales collapses ; p, primordium ; dc, cellule fondatrice du primordium de la nodosité ; gc, cellule géante; o, excroissance; itl, structures semblables à des cordons d'infection, v, vascularisation ; ciz, zone centrale infectée. Les cellules de l'épiderme sont colorées en gris et les cellules corticales en blanc. (modifié d'après Bonaldi *et al.*, 2011).

Contrairement à *A. indica*, *A. afraspera* met en place des nodosités avec ses symbiontes compatible en utilisant la voie des facteurs Nod. Tout d'abord, *A. afraspera* produit deux types

majeurs de flavonoïdes dans ses exsudats : la naringénine et la daidzéine. Ces flavonoïdes induisent la synthèse de facteurs Nod chez *Bradyrhizobium* sp. ORS285 et *B. diazoefficiens* USDA110, respectivement (Renier *et al.*, 2011). Les facteurs Nod d'ORS285 ressemblent beaucoup à ceux produits par USDA110, qui est également capable d'éliciter des nodosités chez *A. afraspera* (Figure 38, Renier *et al.*, 2011).

4.4. La différenciation des bactéroïdes d'*Aeschynomene*

4.4.1. Diversité des morphotypes des bactéroïdes et différenciation chez les symbiontes d'*Aeschynomene*

Les observations des nodosités des deux plantes hôtes *A. indica* et *A. afraspera* ont montré l'existence de morphologies différentes pour la même espèce bactérienne *Bradyrhizobium* sp. ORS285. Chez *A. afraspera*, les bactéroïdes présentent un morphotype allongé (E) comparé aux bactéries résidant dans les excroissances et en vie libre. De même pour *A. indica*, où on peut observer une transition de morphotype indifférencié (U) vers sphérique (S) entre le 4ème et le 5ème jour après l'inoculation. Cette transition passe par une étape d'allongement relatif au cours du 4ème jour, laissant penser que le morphotype sphérique est une forme de différenciation plus poussée dans l'interaction *Bradyrhizobium-Aeschynomene* (Czernic *et al.*, 2015). Il semblerait que les plantes hôtes du groupe d'inoculation III (Nod-indépendantes) induisent une différenciation de leurs bactéroïdes en morphotype S ; et de même pour les plantes du groupe II ayant des bactéroïdes de type E (Nod-dépendantes) (Czernic *et al.*, 2015, Figure 44A).

Comme dans le cas des bactéroïdes présents dans les nodosités d'IRLC, les bactéroïdes d'*Aeschynomene* subissent également un processus d'endoréplication de l'ADN allant jusqu'à 8C pour les bactéroïdes allongés d'*A. afraspera* et 16C pour ceux sphériques d'*A. indica* (Czernic *et al.*, 2015 ; Figure 44B). Un gradient de différenciation quantifiable par le niveau d'endoréplication peut être proposé pour la souche ORS285 avec les morphotypes des bactéroïdes S (16C) > E (8C) > U (1-2C).

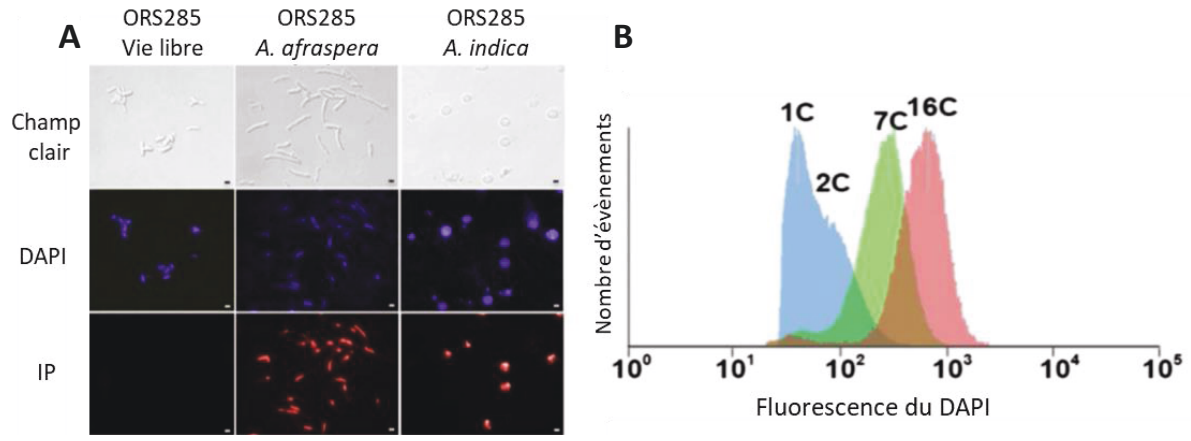


Figure 44: La différenciation cellulaire des bactéroïdes de *Bradyrhizobium* sp. ORS285 en symbiose avec *A. afraspera* et *A. indica*. (A) Comparaisons morphologiques des bactéries en vie libre et des bactéroïdes de la souche ORS285 par microscopie en champ clair et par fluorescence en utilisant le DAPI (intercalant de l'ADN). Les bactéroïdes d'*A. afraspera* et *A. indica* sont respectivement allongés et sphériques. La coloration à l'iodure de propidium (IP) permet d'observer la perméabilisation des membranes des bactéroïdes. (B) Contenu en ADN des bactéries en vie libre (bleu), des bactéroïdes d'*A. afraspera* (vert) et d'*A. indica* (rouge) dont le nombre C correspond au nombre de copies du génome bactérien (d'après Czernic *et al.*, 2015).

4.4.2. Implication de peptides NCR-like d'*Aeschynomene* dans le processus de différenciation

Les déterminants moléculaires du processus de différenciation ont été recherchés en prenant appui sur les découvertes effectuées dans le clade des IRLC. Une analyse transcriptomique comprenant plus de 9000 Expression Sequence Tags (EST) chez les racines et les nodosités d'*A. afraspera* et *A. indica* a permis l'identification de gènes appelés NCR-like, codant des petits peptides riches en cystéines dont la structure est similaire aux NCR des IRLC (Czernic *et al.*, 2015, Figure 45A). Il en existe deux motifs différents, le premier présente un nombre similaire de cystéines à ceux de *Medicago* et le second possède deux cystéines supplémentaires similaires aux défensines de plantes. La phylogénie comparée des NCRs des IRLC et des *Aeschynomene* indique que les séquences d'ADN codant ces peptides ne sont pas homologues, même si leurs origines pointent vers des défensines de plantes dans les deux cas. Ces éléments ont abouti à l'hypothèse que les peptides NCR contrôlant le processus de différenciation des bactéroïdes sont apparus indépendamment au moins deux fois par convergence évolutive. Une protéine homologue à DNF1 de *Medicago* a été identifiée et sa répression par interférence à l'ARN (RNAi) inhibe la différenciation des bactéroïdes d'*A. evenia* comme chez les IRLC (Czernic *et al.*, 2015 ; Figure 45B-C). Cependant, l'activité antimicrobienne et inductrice de

différenciation des NCR-like d'*Aeschynomene* n'a pas été démontrée *in vitro* avec succès sur leurs symbiontes naturels mais uniquement sur la souche *E. meliloti* 1021 (Czernic *et al.*, 2015).

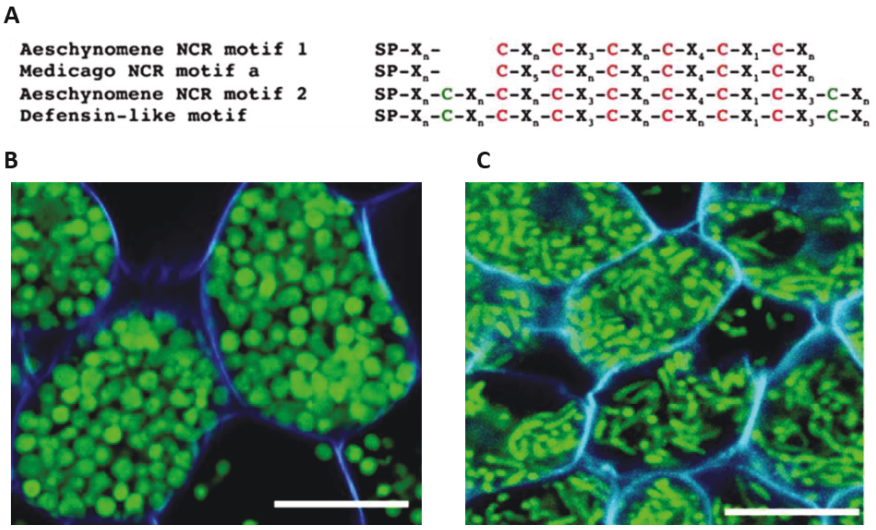


Figure 45: La différenciation des bactéroïdes d'*Aeschynomene* est médiée par des peptides NCR-like. (A) Structure des peptides NCR d'*Aeschynomene* comparée à celles des NCR de *Medicago* et des défensines. (B, C) Images de microscopie confocale montrant la morphologie des bactéroïdes d'*Aeschynoemene evenia* spp. *serrulata* en conditions sauvage (B) ou de répression par RNAi du gène *dnf1* codant une peptidase requise pour la sécrétion des NCR (d'après Czernic *et al.*, 2015).

4.4.3. Identification de BclA, un ABC-transporteur bactérien requis pour la différenciation des bactéroïdes

Au travers d'un crible génétique visant initialement à identifier les déterminants de la symbiose Nod-indépendante entre la souche *Bradyrhizobium* sp. ORS278 et *A. indica*, de nombreux gènes mutés par insertion de transposon ont été révélés comme requis pour la fixation effective de l'azote (Bonaldi *et al.*, 2010). Bien que l'objectif initial n'ait pas été atteint, ce crible a permis de découvrir l'importance des gènes du métabolisme central et des purines pour une fixation effective de l'azote dans le cadre de cette association symbiotique (Bonaldi *et al.*, 2010). De plus, un gène codant un transporteur-ABC putatif présentait des nodosités non fonctionnelles. Ce gène présente une homologie relativement faible (~25%) avec la protéine BacA d'*E. meliloti* et a été nommé *bclA* pour « *Bradyrhizobium bacA-like* ». Son homologue chez la souche ORS285 a été muté et phénotypé sur ses différentes plantes hôtes (Guefrachi *et al.*, 2015, Figure 46). Les bactéroïdes mutants présentent un phénotype non-différencié, un processus d'endoréplication bloqué, et sont quasi incapables de fixer l'azote (Guefrachi *et al.*, 2015). Il y a une mort cellulaire rapide et importante des bactéroïdes dans les nodosités d'*A.*

indica, qui impose un programme de différenciation plus poussé. Enfin, la complémentation du mutant *E. meliloti ΔbacA* avec le gène *bclA* restaure partiellement le phénotype de différenciation, mais pas de fixation. Ce résultat indique que le rôle joué par ces deux protéines est similaire dans la différenciation des bactéroïdes et dans la résistance aux peptides antimicrobiens dont les NCR.

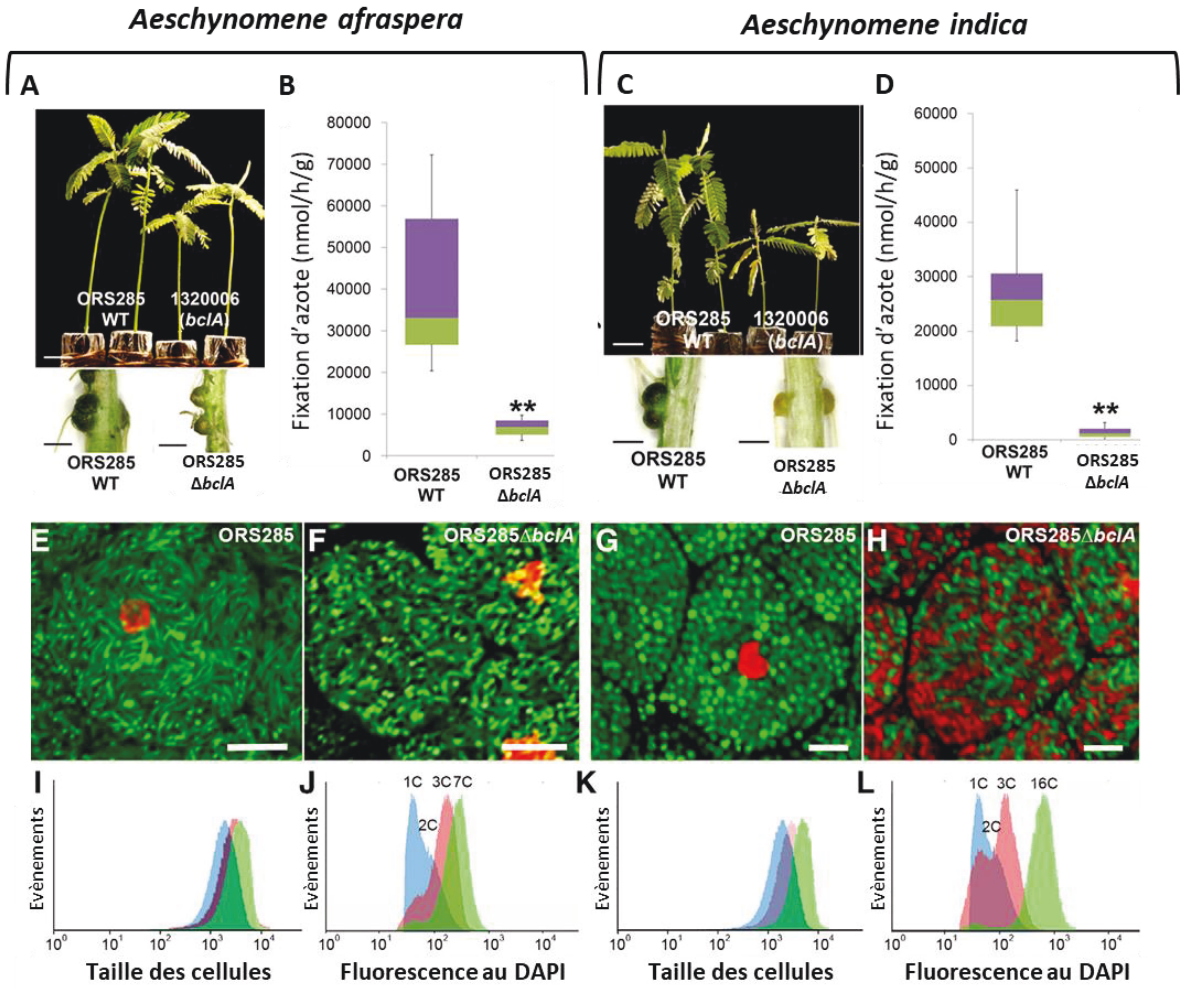


Figure 46: BclA est requis dans le processus de différenciation des bactéroïdes d'ORS285 avec les hôtes *A. afraspera* et *A. indica*. (A, C) Phénotypes des parties aériennes et des nodosités. Les plantes inoculées avec la souche mutée sont plus petites, aux feuilles présentant des signes de chlorose et aux nodosités plus petites, jaunes. (B, D) Mesures de la fixation d'azote par de réduction de l'acétylène comme substrat du complexe enzymatique nitrogénase. Les mutants ne fixent que 20% et 5% de la capacité de fixation de la bactérie sauvage avec *A. afraspera* et *A. indica*, respectivement. (E, F, G, H) Images de microscopie confocale montrant l'absence de différenciation des bactéroïdes pour les deux conditions mutantes, et même une mort cellulaire très importante des bactéroïdes d'*A. indica* (H) à 14 dpi. (I ,J , K, L) Expériences de cytométrie en flux mesurant la taille des cellules (I, K) et leur niveau de ploïdie (J, L). La ploïdie des bactéroïdes mutants ne dépasse pas 3C, là où les bactéroïdes sauvages atteignent 8C et 16C chez *A. afraspera* et *A. indica*, respectivement (modifié d'après Guefrachi *et al.*, 2015).

5. L'apport des données -omiques et notion(s) d'efficacité symbiotique

5.1. Comparaison des différentes approches -omiques

L'intérêt des approches dites « -omiques » consiste à appréhender la complexité d'un système biologique dans son ensemble, en utilisant des méthodologies les moins restrictives sur le plan descriptif. Cela signifie que ce sont des techniques à haut débit où un grand nombre de variables sont analysées en même temps. Les types d'analyses -omiques considérées ici sont la transcriptomique (expression des gènes et leur régulation), la protéomique (analyse des protéines présentes) et la métabolomique (étude des métabolites produits). Ces approches permettent d'obtenir de très nombreuses informations sur la réponse cellulaire et/ou tissulaire à différentes conditions d'intérêt. Ces approches sont sans *a priori*, c'est-à-dire ne ciblant pas un gène, un transcrit ou une protéine en particulier comme c'est le cas en utilisant respectivement les techniques de PCR, de RT-qPCR et d'immunomarquage. Cela peut aussi signifier sans *a priori* sur le rôle possible de gène/métabolite, au contraire d'un mutant issu d'un crible génétique chosi pour son phénotype observé. Ces méthodes peuvent mettre en évidence des gènes appartenant à des régulons ou des transcrits/protéines/métabolites différentiellement exprimés/accumulé(e)s par contraste entre deux ou plusieurs conditions.

Les approches transcriptomiques sont celles qui ont connu la plus grande avancée technologique ces dernières années. Les premières technologies de macroarrays et microarrays se basaient sur l'hybridation de sondes spécifiques de transcrits déjà connus et en nombre limité. Cela nécessitait donc des connaissances *a priori* sur les séquences nucléotidiques appartenant à l'organisme étudié. L'émergence des techniques de séquençage de nouvelle génération a permis le saut technologique du RNA-seq. Cette méthode a de nombreux avantages. Par exemple, il n'est pas nécessaire de connaitre *a priori* des séquences nucléotidiques et il est possible d'identifier et quantifier tous les ARNm de l'échantillon si la profondeur de séquençage est suffisante (Manzoni *et al.*, 2016).

Cependant chaque technologie -omique ne permet pas de considérer le même nombre de variables ni d'atteindre le même niveau d'interprétation. En effet, une approche transcriptomique depuis l'avènement du RNA-Seq, est capable de donner le niveau d'expression de la majorité des gènes du tissu cellulaire étudié. Cependant, une expression importante n'informe pas nécessairement sur ce qu'il se passe réellement au niveau phénotypique du fait de possibles régulations post-transcriptionnelles et post-traductionnelles. De plus, l'expression dudit gène n'est pas forcément liée à un phénotype observable si ce

dernier est muté. Ce type d'approche dispose donc d'une couverture large mais les interprétations se limitent tout de même à décrire « ce qui peut arriver » (Figure 47). Au contraire, les approches de métabolomique ne considèrent qu'un nombre réduit de molécules mais représentent plus directement la réalité du fonctionnement cellulaire. Les approches protéomiques sans marquage sont quant à elles entre ces deux extrêmes en étant plus résolutive en termes d'interprétations que la transcriptomique car on quantifie directement « ce qui peut faire arriver », notamment les enzymes. Cependant, un défi majeur serait de couvrir la gamme dynamique de l'expression de ces protéines car les protéines membranaires et les protéines minoritaires (comme les facteurs de transcription) sont plus difficiles à extraire et donc à détecter au sein d'un échantillon biologique. Actuellement, il est plus aisé d'identifier et de quantifier l'expression des gènes codant ces protéines par des approches transcriptomiques.

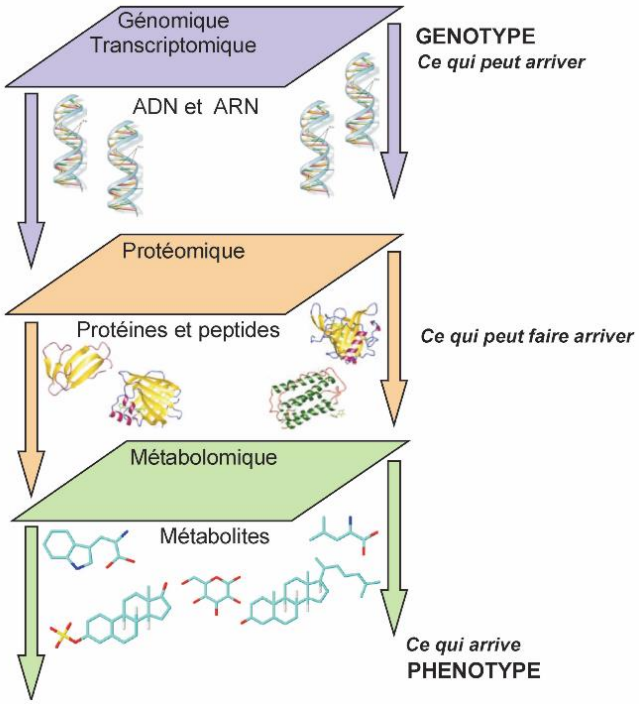


Figure 47: Représentation schématique des trois niveaux de caractérisation des systèmes biologiques par des approches -omiques (d'après : http://www.ipubli.inserm.fr/bitstream/handle/10608/222/?sequence=30).

5.2. Apport des approches -omiques dans la compréhension des mécanismes symbiotiques entre rhizobia et légumineuses

Les approches -omiques ont identifié un grand nombre de mécanismes moléculaires impliqués dans l'établissement et le fonctionnement des symbioses entre les rhizobia et les légumineuses, ainsi que des interactions plantes / micro-organismes en général. Les études transcriptomiques et protéomiques ont d'abord été effectuées sur les partenaires bactériens et principalement sur *E. meliloti* et *B. diazoefficiens* du fait de l'obtention plus aisée de leur génome dès le début des années 2000 (Galibert *et al.*, 2001 ; Kaneko *et al.*, 2002). A titre de comparaison, la séquence génomique de *Medicago truncatula* n'a été publié que 10 ans plus tard (Young *et al.*, 2011). Ne seront mentionnées ici que les approches considérant le partenaire bactérien. La première étude en date concerne la souche *E. meliloti* 1021 en utilisant des macroarrays (Ampe *et al.*, 2003). Depuis lors, les études -omiques ont apporté des connaissances plus vastes sur les aspects des plus précoces aux plus tardifs du processus symbiotique entre les rhizobia et les légumineuses (Figure 48A). Leurs apports ont été résumés de façon quasi-exhaustive dans des revues récentes (Jiménez-Guerrero *et al.*, 2018 ; diCenzo *et al.*, 2018). Les conditions comparées les plus employées vont être résumées ici au travers de quelques exemples choisis (Figure 48B).

L'environnement des bactéroïdes comme symbiontes intracellulaires de végétaux est très complexe. Cela rend difficile l'interprétation de modifications de l'expression des gènes bactériens, contrairement à des milieux liquides dont les conditions physico-chimiques sont contrôlées. La communauté scientifique a donc cherché à décomposer le processus symbiotique en mimant de façon modulaire un seul paramètre à la fois parmi des conditions physico-chimiques déjà connues de la symbiose. L'effet de flavonoïde(s) sur la cascade de signalisation NodD en culture libre a souvent été étudié pour établir quelles sont les réponses cellulaires lors des étapes précoces de l'infection (Ampe *et al.*, 2003, Barnett *et al.*, 2004 ; Capela *et al.*, 2005 ; Pérez-Montaño *et al.* 2016a ; 2016b). Des études aux résultats similaires ont été effectuées en protéomique, mais également la détermination du sécrétome par le T3SS des rhizobia induite par des flavonoïdes (Hempel *et al.*, 2009 ; Okazaki *et al.*, 2010). Pour mimer les étapes tardives de l'infection, l'effet d'une culture en microaérobie limitant ainsi la concentration en O_2 et déclenchant l'expression des gènes symbiotiques a été utilisé. Ces expériences ont mené à préciser l'implication des régulateurs connus comme FixLJ et NifA précédemment décrits et leurs cibles potentielles (Ampe *et al.*, 2003 ; Becker *et al.*, 2004 ; Pessi *et al.*, 2007). Depuis la

découverte de l'effet des NCR sur le processus de différenciation *in planta*, de nouvelles conditions ont été testés *in vitro*. Les effets des peptides NCR247 et NCR335 sur le transcriptome de *E. meliloti* ont récemment été testés (Tiricz *et al.*, 2013, Penterman *et al.*, 2014).

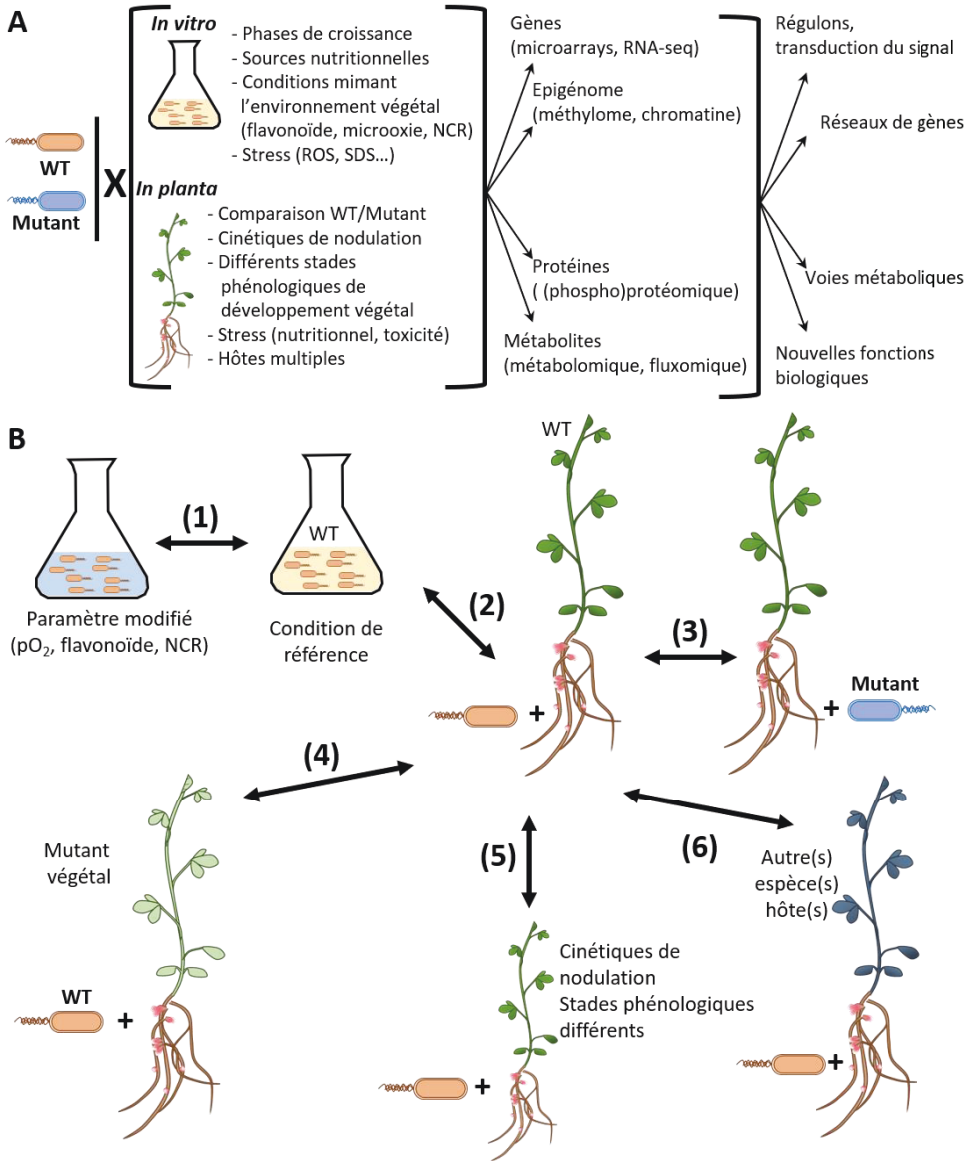


Figure 48: Résumés des systèmes expérimentaux fréquemment utilisés lors d'approches –omiques appliquées aux bactéries symbiotiques de légumineuses. (A) Ensemble de combinaisons possibles de comparaisons de conditions utilisant diverses méthodes -omiques. (B) Sélection de 6 types de systèmes expérimentaux développés ici. (1) Comparaison en milieu liquide entre une culture de référence et une condition modifiée, permettant de définir des régulons. (2) Comparaison entre les bactéroïdes sauvages et une culture libre pour déterminer les facteurs considérés en conditions symbiotiques. (3) Comparaison entre deux conditions de bactéroïdes au génotype sauvage ou mutant pour observer l'effet de la mutation sur la réponse cellulaire et moléculaire du mutant (4) Utilisation de mutants végétaux pour observer la réponse du partenaire bactérien comparé à la plante sauvage. (5) Utilisation de différents stades phénologiques afin de suivre l'évolution temporelle de la physiologie des bactéroïdes. (6) Comparaisons entre divers hôtes infectés par la même souche bactérienne pour d'accéder aux spécificités et à l'adaptation de la souche bactérienne pour chaque association. (Image plante : crédit Quentin Nicoud)

Les bactéroïdes fixateurs d'azote dans les nodosités ont été comparés à une condition de référence en culture libre aérobie par des approches transcriptomiques (Ampe *et al.*, 2003 ; Pessi *et al.*, 2007 ; Karunakaran *et al.*, 2009) et métabolomiques (Vauclare *et al.*, 2013). Le consensus général ressortant de toutes ces études est l'observation de changements drastiques, où l'on observe par exemple des milliers de gènes différentiellement exprimés quand on compare ces deux conditions en transcriptomique. Un grand nombre de gènes sont réprimés en conditions symbiotiques et les gènes impliqués dans la fixation d'azote sont très induits, signe de la spécialisation métabolique extrême que subissent les bactéroïdes.

Des mutants bactériens impactés dans le processus symbiotique ont également été abondamment utilisés car ils facilitent l'identification des régulons de régulateurs transcriptionnels comme c'est le cas des mutants bactérien *nifA* et *rpoN1/2* chez *B. diazoefficiens* USDA110 (Hauser *et al.*, 2007 ; Pessi *et al.*, 2007 ; Lardi *et al.*, 2016). De même, des mutants végétaux présentant un blocage à diverses étapes de la mise en place des nodosités ont été utilisés comme c'est le cas des mutants *dnf* de *M. truncatula* (Lang et Long, 2015).

Dans le but d'observer les variations temporelles d'expression des gènes dans la nodosité, la communauté a également effectué des analyses avec des nodosités d'âges variés (Ampe *et al.*, 2003 ; Capela *et al.*, 2006). Récemment, les motifs d'expression spatiale des gènes des nodosités ont été révélés en couplant une analyse transcriptomique à une microdissection laser des zones des nodosités (Roux *et al.*, 2014).

Enfin, différents hôtes végétaux d'un même symbionte ont été étudiés par ces approches pour définir les adaptations spécifiques des bactéries à un hôte donné (Karunakaran *et al.*, 2009 ; Koch *et al.*, 2010 ; Lardi *et al.*, 2016).

5.3. L'efficacité symbiotique dans les associations Rhizobium-Légumineuses

5.3.1. Des questions de définition, d'échelle et de point de vue

Nous avons donc vu l'intérêt de l'utilisation des technologies -omiques qui mènent à une meilleure compréhension du processus symbiotique, ouvrant ainsi les portes d'une possible optimisation de ce processus dans un deuxième temps. Mieux comprendre quelles sont les signatures moléculaires des variations de performances dans les associations symbiotiques permettra alors de maximiser l'efficacité de ces dernières. En effet, il a depuis longtemps été

compris que diverses associations de génotypes des plantes hôtes et des rhizobia pouvaient donner lieu à de nombreuses différences en termes de nombre de nodosités, de croissance végétative de l'hôte, de persistance des bactéries dans le sol, ou de rendement en graines (Thilakarathna et Raizada, 2017). Le choix des associations en fonction de leur efficacité symbiotique est un sujet d'intérêt majeur. Seulement, en analysant la littérature concernant l'évaluation de l'efficacité symbiotique des interactions rhizobium-légumineuses, on peut observer que le concept est rarement défini et que la plupart des études se limitent à quantifier des traits phénotypiques de l'hôte végétal (Gourion *et al.*, en préparation, Figure 49). Du fait des conditions de laboratoire, cela se limite également à comparer la croissance végétative de l'hôte, sans aller jusqu'au rendement en graines qui est plus intéresssant pour les agronomes, qui mesure mieux la valeur sélective de l'hôte. Ainsi, il est sans doute nécessaire d'établir une définition précise de ce concept et de discuter de l'intérêt des paramètres végétaux et bactériens qui contribuent à ce concept global d'efficacité symbiotique. La notion même de symbiose, où les deux partenaires peuvent tirer un bénéfice dans leur interaction, mène à un équilibre dynamique quant aux bénéfices obtenus pour l'un ou l'autre. Sans même aller jusqu'à considérer une course aux armements, l'émergence de génotypes qui maximisent leur valeur sélective au détriment de celle de leur partenaire est difficilement évitable. Si on se place uniquement dans la perspective d'un des deux partenaires symbiotiques, les paramètres choisis pourraient alors mener à des conclusions opposées.

L'efficacité d'un système peut être vue comme la quantité d'énergie qui doit être allouée (un apport) pour effectuer une action particulière (e.g une réaction enzymatique avec une production quantifiable). Cette définition peut être valide d'un point de vue biochimique avec par exemple l'équation bilan du fonctionnement du complexe nitrogénase. Mais elle devient plus restrictive quand il s'agit de prendre en compte un contexte plus large d'interaction symbiotique qui inclut d'autres paramètres pré et post-infectieux. En effet, la compétitivité entre les bactéries dans le sol et pendant le processus d'infection, la qualité du dialogue moléculaire entre les partenaires ainsi que le déroulement de l'organogénèse et son coût vont avoir un impact sur l'établissement et le fonctionnement global de l'interaction symbiotique. Une fois les bactéroïdes formés, la fixation d'azote varie en fonction des souches, et le résultat de l'interaction devrait être considéré pour les deux partenaires, en analysant à la fois le rendement en graines des plantes et la valeur sélective bactérienne en mesurant le nombre (la persistance) des rhizobia dans le sol après la culture (Jensen et Sørensen, 1987). En effet, même après une symbiose effective pour la plante, rien ne garantit la survie des rhizobia, qui dépend de leur

compétitivité face aux autres bactéries du sol jusqu'à la prochaine culture. On ne considérera pas ici le point de vue capitaliste, selon lequel il serait plus rentable de vendre chaque année un inoculum de rhizobia associé ou enrobées aux graines utilisées pour la culture. Le meilleur résultat pour les rendements à l'échelle planétaire serait d'avoir un panel de rhizobia efficaces et persistantes dans le sol dans un maximum de biomes où la culture de légumineuses est possible. Le problème étant que la valeur sélective des bactéries n'a que rarement été mesurée, en dehors d'études des sanctions que peut appliquer l'hôte à de potentiels tricheurs bactériens (Kiers *et al.*, 2003). En effet, dans une condition de triche artificielle où l'azote de l'air est remplacé par de l'argon, rendant la fixation d'azote impossible, les rhizobia ayant colonisé les nodosités présentent un succès reproducteur diminué de moitié (Kiers *et al.*, 2003).

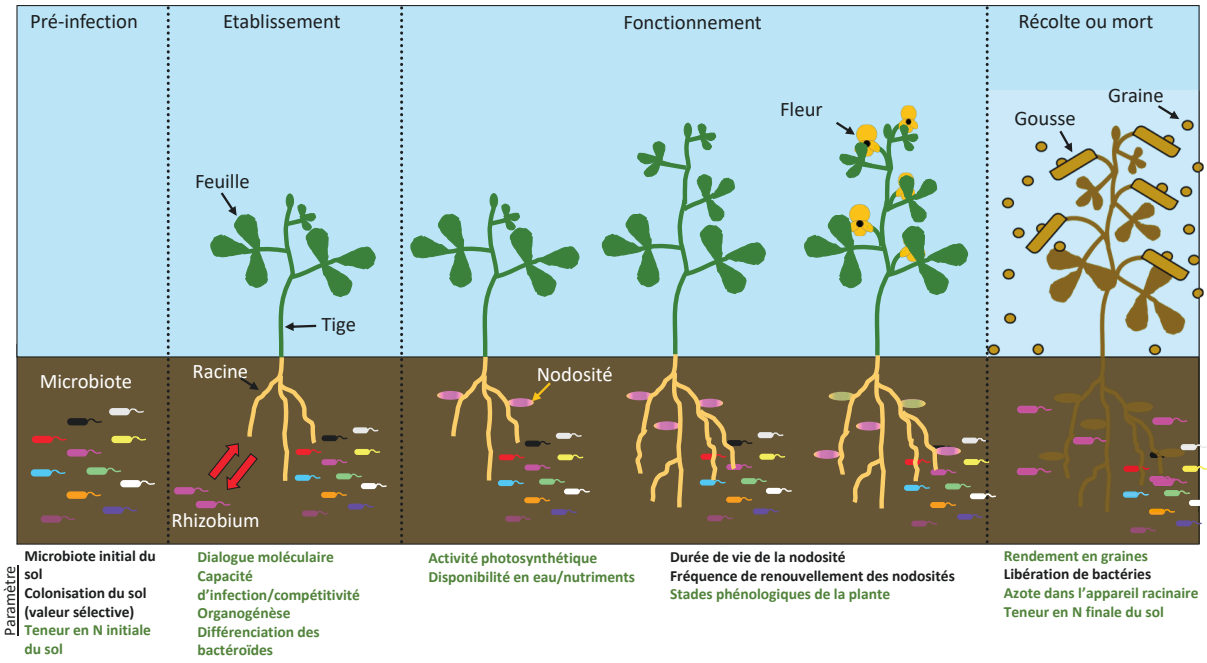


Figure 49: Cycle de vie de l'interaction symbiotique entre rhizobia et une légumineuse annuelle, et paramètres à quantifier pour mesurer l'efficacité et le bénéfice de l'interaction pour les deux partenaires. Les paramètres en vert sont ceux à étudier pour maximiser la valeur sélective de la plante hôte seulement (modifié d'après Gourion *et al.*, en préparation)

5.3.2. Valeur sélective bactérienne et différenciation des bactéroïdes

Bien que peu étudiée actuellement, la quantification de la valeur sélective bactérienne est essentielle pour comprendre le maintien de ce type d'interaction symbiotique qui existe depuis plus de 60 Ma (Remigi *et al.*, 2016). Cette quantification est d'autant plus importante à mesurer dans le cas des IRLC et leurs peptides NCR qui induisent une différenciation terminale (Alunni

et Gourion, 2016). Leur existence depuis plus de 40 Ma (Sprent *et al.*, 2013) implique que cette différenciation terminale n'empêche pas la symbiose de perdurer dans le temps. L'hypothèse actuelle serait que l'augmentation de la valeur sélective des bactéries colonisant des nodosités d'IRLC, supposés clonales, se fasse via leur survie et leur reproduction dans les cordons d'infection où elles restent non différenciées et sont libérées après la sénescence de la nodosité. Cela laisse supposer une forme de sélection de parentèle bactérienne (*kin selection*) où la population des bactéroïdes différenciés n'aura pas de descendance au bénéfice de celle des bactéries (clonales) dans les cordons d'infection. D'autre part, l'émergence attendue de stratégies d'échappement à cette sélection a d'ailleurs déjà été observée. Des souches de *E. meliloti* possédant une métallopeptidase HrrP (*Host range restriction peptidase*) présente sur un plasmide accessoire ont un phénotype Fix^- dépendant du génotype de l'hôte végétal (Price *et al.*, 2015). La peptidase HrrP est capable de cliver des peptides NCR et induit une sénescence accélérée des nodosités. Les bactéroïdes semblent subir une forme de dédifférenciation de la zone centrale à la zone proximale de la racine et ont un meilleur taux de survie quand extraits des nodosités (Price *et al.*, 2015). Dans cet exemple, on a un cas d'augmentation de la valeur sélective de l'endosymbionte au détriment de celle de l'hôte, qui n'en tire qu'un bénéfice réduit.

5.3.3. Différenciation et efficacité symbiotique du point de vue végétal

A l'instar de *Bradyrhizobium* sp. ORS285 qui est un symbionte de plusieurs plantes qui imposent différents programmes de différenciation, d'autres exemples intéressants de souches bactériennes capables de noduler différentes plantes avec des morphotypes variés ont été étudiés (Oono et Denison, 2010). L'étude porte sur deux souches. *R. leguminosarum* A34 présente des morphotypes U et E en symbiose avec *Phaseolus vulgaris* (haricot) et *Pisum sativum* (pois), respectivement. *Bradyrhizobium* 32H1 a des morphotypes U et S en symbiose avec *Vigna unguiculata* (niébé) et *Arachis hypogaea* (arachide), respectivement. Les auteurs ont essayé de mesurer l'efficacité de la symbiose des bactéroïdes différenciés et l'ont comparée à celle de bactéroïdes indifférenciés. L'estimateur utilisé était le ratio H_2/CO_2. Ce ratio mesure ainsi le taux de fixation d'azote par celui du H_2, son autre produit, par rapport au taux de respiration dégageant du CO_2. Leur conclusion a été que les bactéroïdes différenciés semblent être plus efficaces en termes de fixation d'azote que les bactéroïdes non différenciés (Oono et Denison, 2010 ; Figure 50). Si les différences entre les morphotypes différenciés et non différenciés d'une même souche sur diverses plantes hôtes ont déjà été observés, il n'est

mentionné nulle part dans la littérature de question visant à comparer des associations plante-symbionte où il y a dans les deux cas un processus de différenciation.

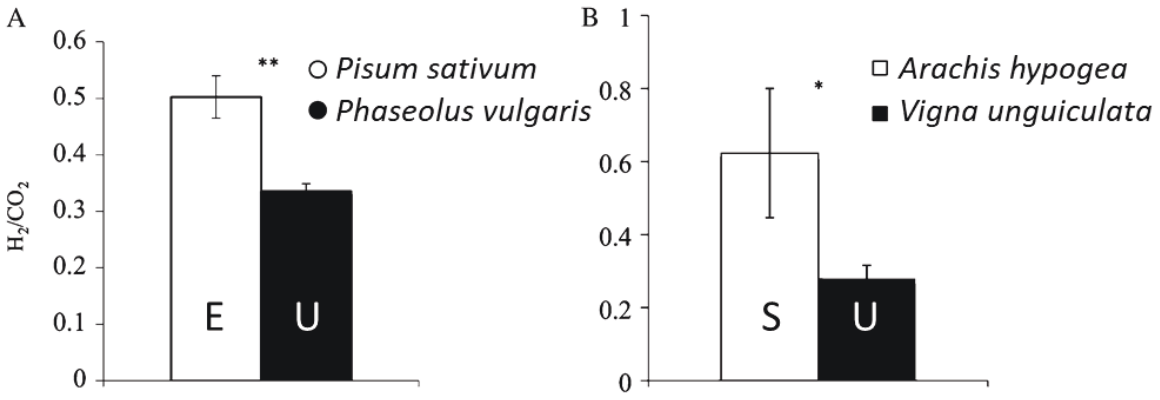


Figure 50: Comparaison de l'efficacité symbiotique de bactéroïdes différenciés et non différenciés. (A) Mesures de la fixation d'azote de *R. leguminosarum* A34 présentant des morphotypes E et U en symbiose avec *P. sativum* et *P. vulgaris* respectivement. (B) Mesures de la fixation d'azote de *Bradyrhizobium* sp. 32H1 présentant des morphotypes S et U en symbiose avec *A. hypogaea* et *V. unguiculata* respectivement. Dans les deux cas, les bactéroïdes différenciés présentent une fixation d'azote plus importante que les non différenciés (modifié d'après Oono et Denison, 2010).

5.4. Problématiques et objectifs du projet de thèse

C'est dans le cadre de ces problématiques multiples soulevées par le modèle *Aeschynomene-Bradyrhizobium* précédemment décrit que s'inscrit mon projet de thèse. Avant mon arrivée au laboratoire, la possible corrélation entre différenciation et efficacité symbiotique a été explorée avec plus de précision en partant des associations symbiotiques entre les genres *Medicago* et *Ensifer*. J'ai pris part à cette étude suggérant une corrélation positive entre l'efficacité de la symbiose pour le partenaire végétal et le niveau de différenciation des bactéroïdes mesurée par cytométrie en flux (Kazmierczak *et al.*, 2017 ; Annexe 1). La poursuite de ce questionnement s'applique également très bien au système *Bradyrhizobium-Aeschynomene* comme une même souche présente deux morphotypes de bactéroïdes dont les degrés de différenciation sont très distincts en fonction de l'hôte végétal. La physiologie des symbiontes du genre *Bradyrhizobium* dans les nodosités d'*Aeschynomene* pendant une symbiose fonctionnelle n'a pas encore été décrite par des approches globales sans *a priori*. De plus, les connaissances actuelles sont limitées concernant les mécanismes moléculaires bactériens d'adaptation des bradyrhizobia à différents hôtes végétaux imposant divers degrés de différenciation (non différenciés, allongés ou sphériques). Par ailleurs des souches mutantes pour le transporteur BclA sont disponibles,

et le degré de nécessité de cette protéine pour la fixation d'azote dépend de l'hôte végétal (Guefrachi *et al.*, 2015 ; Barrière *et al.*, 2017 ; Annexe 2).

Ainsi, les questions biologiques auxquelles nous avons tenté de répondre au cours de cette thèse sont les suivantes :

- Existe-t-il une corrélation entre l'efficacité de la symbiose et le processus de différenciation dans les associations symbiotiques entre les genres *Bradyrhizobium* et *Aeschynomene* ?
- Quels sont les gènes/protéines des symbiontes impliqués dans le processus normal de fixation d'azote en association avec les plantes du genre *Aeschynomene* ?
- Quelles sont les signatures moléculaires de la différenciation des bactéroïdes dans ces associations symbiotiques ? Sont-elles similaires à celles des IRLC ?
- Quels sont les facteurs bactériens d'adaptations spécifiques de divers hôtes végétaux imposant des programmes de différenciation distincts ?

Nous avons donc décidé de comparer l'efficacité symbiotique des morphotypes E et S à partir de souches isolées de nodosités d'*Aeschynomene* capables de noduler les groupes d'inoculation II et III, comme la souche *Bradyrhizobium* sp. ORS285 (Alazard, 1985 ; Molouba *et al.*, 1999). Des mesures de masses de plantes en association avec diverses souches de *Bradyrhizobium* ont été effectuées. Parallèlement, des mesures de ploïdie des bactéroïdes par cytométrie en flux ont permis de quantifier la corrélation possible existant entre l'efficacité de la symbiose et le processus de différenciation des bactéroïdes (Chapitre 1).

La souche *Bradyrhizobium* sp. ORS285 devient un modèle émergent avec de nombreuses spécificités pouvant compléter des connaissances relatives à la diversité des processus symbiotiques et même au-delà. Comme mentionné précédemment et contrairement à une majorité de rhizobia, elle est capable de photosynthèse et de noduler des *Aeschynomene* des groupes d'inoculation II (nodulation Nod-dépendante) et III (nodulation Nod-indépendante). Cette souche a également été séquencée et annotée de façon experte (Mornico *et al.*, 2012 ; Gully *et al.*, 2017). Les bactéroïdes de la souche ORS285 ont été choisis pour comprendre leur physiologie en analysant les transcriptomes en symbiose avec *A. indica* ou *A. afraspera* (Figure 51A). Les résultats observés concernant l'efficacité symbiotique dans le chapitre 1 ont été consolidés pour cette seule souche par des analyses élémentaires et fluxomiques. Des analyses fonctionnelles de gènes candidats différentiellement exprimés par rapport à une culture de référence ou spécifiques d'un des deux hôtes a été menée (Chapitre 2).

Une comparaison combinant transcriptomique et métabolomique entre les génotypes sauvages et Δ*bclA* d'ORS285 sur les mêmes plantes hôtes a également été effectuée. Elle permet ainsi de comparer la physiologie des bactéroïdes fonctionnels et non fonctionnels, et plus largement de nodosités où peuvent avoir lieu une symbiose fonctionnelle ou non (Chapitre 3).

Cependant, à notre connaissance, il n'existe pas de plante hôte n'imposant pas de programme de différenciation des bactéroïdes et capable de former des nodosités fonctionnelles avec la souche ORS285. Du fait de sa grande importance agronomique, le soja aurait été un excellent candidat comme plante hôte induisant un morphotype U. Il peut être nodulé par USDA110 qui partage des facteurs Nod très similaires à ceux de la souche ORS285 (Renier *et al.*, 2011), et il existe une publication montrant une symbiose effective entre le soja et la souche ORS285 (Giraud *et al.*, 2013). Ces résultats n'ont malheureusement jamais pu être reproduits. Cependant, la plante hôte *A. afraspera* peut également être nodulée par *B. diazoefficiens* USDA110. Nous avons donc comparé les transcriptomes et les protéomes de la souche USDA110 en symbiose avec *A. afraspera* et le soja (Figure 51B). Notre but était d'évaluer l'adaptation de cette bactérie à un hôte qui impose un programme de différenciation à ses symbiontes en sécrétant des peptides NCR. Nous avons également effectué un métabolome de nodosités complètes et des analyses fonctionnelles de gènes d'intérêt ressortant de cette intégration de jeux de données-omiques. En supplément, cette étude comprend aussi une comparaison de l'efficacité symbiotique de la survie et de la perméabilité des bactéroïdes d'USDA110 entre *A. afraspera* et *G. max* (Chapitre 4).

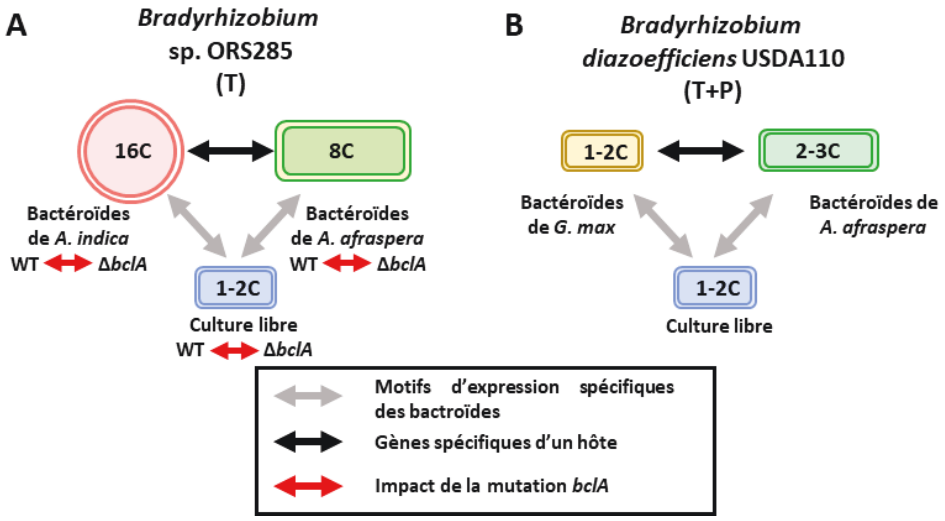
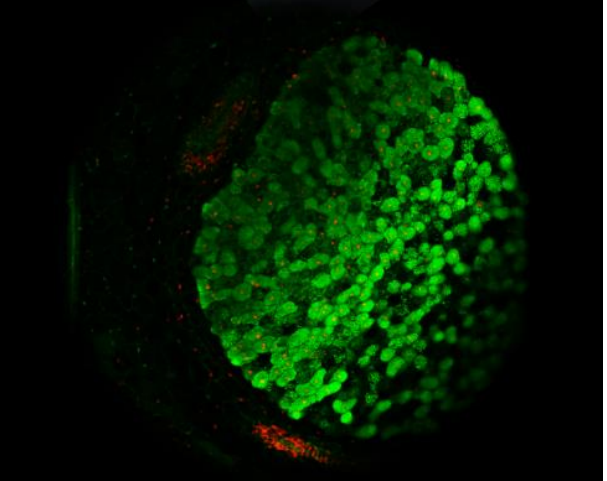
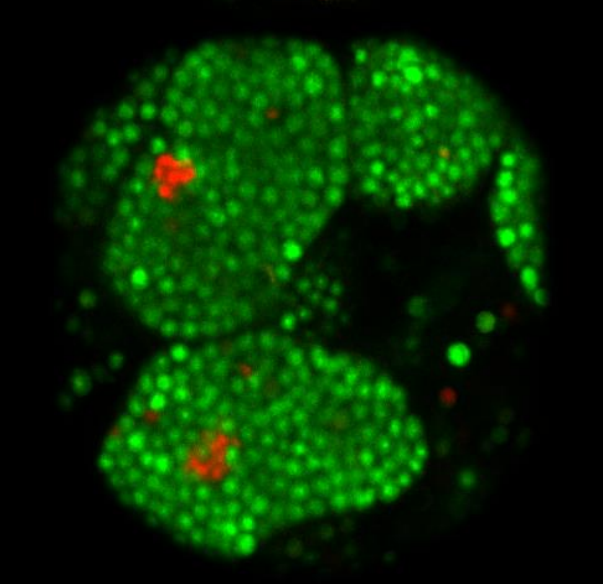


Figure 51: Systèmes expérimentaux et comparaisons effectuées au cours des analyses transcriptomiques (T) et protéomiques (P). Comparaisons de conditions impliquant les souche ORS285 (A) et USDA110 (B).





Partie II : Résultats

Partie II : Résultats

1. Symbiotic efficiency of spherical and elongated bacteroids in the *Aeschynomene-Bradyrhizobium* symbiosis

Florian Lamouche, Nolwenn Bonadé-Bottino, Peter Mergaert et Benoit Alunni

Avant-propos

Cette étude a été précédée par la découverte et la quantification de peptides NCR-like chez *Aeschynomene* qui induisent un processus de différenciation des bactéroïdes déjà observé depuis longtemps chez les Dalbergioïdes (Sen et Weaver, 1984 ; Van de Velde *et al.*, 2010 ; Czernic *et al.*, 2015). En fonction de l'espèce végétale, les *Aeschynomene* imposent des modifications morphologiques à leurs microsymbiontes, leur donnant des morphotypes allongés (E) ou sphériques (S) (Czernic *et al.*, 2015).

Parallèlement, des chercheurs se sont focalisés sur une recherche exhaustive, au sein du clade des légumineuses, de traits phénotypiques tels que le caractère déterminé ou indéterminé des nodosités, et l'existence d'un processus de différenciation des bactéroïdes ou non (Oono *et al.*, 2010). Le bénéfice potentiel conféré par la différenciation des bactéroïdes a été investigué dans l'étude précédemment décrite en fin d'introduction (Figure 50), dont la conclusion était que les plantes hôtes imposant un programme de différenciation des bactéroïdes en tirent un bénéfice plus important (Oono et Denison, 2010).

C'est en intégrant ces deux axes de recherche que nous nous sommes interrogés sur la quantification et la comparaison du bénéfice symbiotique pour la plante en fonction de si elle impose un programme de différenciation donnant des bactéroïdes de type S ou E. Dans ce travail, nous avons procédé à des récoltes des différents organes végétaux de 3 espèces différentes d'*Aeschynomene* 15 jours après l'inoculation par 6 différentes souches bactériennes isolées de nodosités d'*Aeschynomene* spp. (Molouba *et al.*, 1999). J'ai réalisé l'ensemble des expériences présentées avec l'aide de Stacy Simoes (tagiaire L2), Nolwenn Bonadé-Bottino (stagiaire L3) et Corine Foucault.

Ce travail a été publié dans la revue *Frontiers in Plant Science* en 2019. DOI : https://doi.org/10.3389/fpls.2019.00377

frontiers in Plant Science

ORIGINAL RESEARCH
published: 02 April 2019
doi: 10.3389/fpls.2019.00377

Symbiotic Efficiency of Spherical and Elongated Bacteroids in the *Aeschynomene-Bradyrhizobium* Symbiosis

*Florian Lamouche, Nolwenn Bonadé-Bottino, Peter Mergaert and Benoit Alunni**

Institute for Integrative Biology of the Cell, UMR 9198, CNRS/Université Paris-Sud/CEA, Université Paris Saclay, Gif-sur-Yvette, France

The legume-rhizobium symbiosis is a major supplier of fixed nitrogen in the biosphere and constitutes a key step of the nitrogen biogeochemical cycle. In some legume species belonging to the Inverted Repeat Lacking Clade (IRLC) and the Dalbergioids, the differentiation of rhizobia into intracellular nitrogen-fixing bacteroids is terminal and involves pronounced cell enlargement and genome endoreduplication, in addition to a strong loss of viability. In the *Medicago truncatula-Sinorhizobium* spp. system, the extent of bacteroid differentiation correlates with the level of symbiotic efficiency. Here, we used different physiological measurements to compare the symbiotic efficiency of photosynthetic bradyrhizobia in different *Aeschynomene* spp. (Dalbergioids) hosts inducing different bacteroid morphotypes associated with increasing ploidy levels. The strongly differentiated spherical bacteroids were more efficient than the less strongly differentiated elongated ones, providing a higher mass gain to their hosts. However, symbiotic efficiency is not solely correlated with the extent of bacteroid differentiation especially in spherical bacteroid-inducing plants, suggesting the existence of other factors controlling symbiotic efficiency.

Keywords: legume-rhizobium symbiosis, bacteroid differentiation, symbiotic efficiency, flow cytometry, *Aeschynomene*, *Bradyrhizobium*

OPEN ACCESS

Edited by:
Lorena Carro, Newcastle University, United Kingdom

Reviewed by:
Erik Limpens, Wageningen University and Research, Netherlands
Muthu Venkateshwaran, University of Wisconsin–Platteville, United States

***Correspondence:**
Benoit Alunni
benoit.alunni@i2bc.paris-saclay.fr

Specialty section:
This article was submitted to Plant Microbe Interactions, a section of the journal Frontiers in Plant Science

Received: *28 September 2018*
Accepted: *12 March 2019*
Published: *02 April 2019*

Citation:
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. doi: 10.3389/fpls.2019.00377

INTRODUCTION

Nitrogen-fixing symbioses constitute a major process in the nitrogen biogeochemical cycle, accounting for the main input of fixed nitrogen in terrestrial ecosystems (Gruber and Galloway, 2008). The interaction between rhizobia and their legume hosts leads to the formation of root or stem nodules that house the bacterial symbiont (Sprent et al., 2017). After mutual recognition of the partners through a molecular dialogue, a coordinated program of organogenesis and infection results in a mature nodule where plant symbiotic cells are filled with intracellular bacteria that differentiate into nitrogen-fixing bacteroids (Kondorosi et al., 2013).

Bacteroids do not always exhibit the same morphological features as their kin in the soil. In some legume clades such as the Inverted Repeat Lacking Clade (IRLC) to which *Medicago truncatula* belongs, bacteroids undergo a strong cell enlargement, associated to an increased membrane permeability. During this process, elongated bacteroids become polyploid and their capacity to resume growth is compromised. Their differentiation is thus considered to be terminal

(Mergaert et al., 2006). These bacteroids are referred to as E-type bacteroids (i.e., E for elongated). Oppositely, in most legume clades like the ones comprising *Lotus japonicus* or common bean (*Phaseolus vulgaris*), bacteroids remain similar in size and shape to the free-living rhizobia. They also retain full capacity to resume growth when they are extracted from nodules (Mergaert et al., 2006). They are thus not terminally differentiated and are called U-type bacteroids (i.e., U for unmodified).

The phenomenon of terminal bacteroid differentiation is controlled by the host plant through the massive production of a large array of nodule-specific cysteine-rich (NCR) antimicrobial peptides (Mergaert et al., 2003; Van de Velde et al., 2010). In *M. truncatula*, more than 600 NCR genes are encoded in the genome and several of them are clustered into restricted genomic locations where gene duplication and exchanges of gene pieces between paralogs result in a dynamic growth of this gene repertoire (Alunni et al., 2007; Kondorosi et al., 2013; Pecrix et al., 2018). Remarkably, despite the large number of NCR genes suggesting *a priori* an extensive redundancy, some of the NCR genes in *M. truncatula* are essential and their mutation results in bacteroid defects and a non-functional symbiosis (Horváth et al., 2015; Kim et al., 2015; Wang et al., 2017; Yang et al., 2017). Within the IRLC, the extent of bacteroid differentiation varies depending on the host plant and is correlated with the number of NCR genes in its genome, especially with the number of cationic ones (Montiel et al., 2017).

At the functional level, it has been shown that cationic NCRs display a pore-forming activity that induces membrane permeability and ultimately microbial death (Van de Velde et al., 2010; Mikuláss et al., 2016). NCR peptides also rewire the bacterial transcriptome, showing features of a membrane stress response (Tiricz et al., 2013). The bacterial peptide transporter BacA is critical to protect the bacteroids against this antimicrobial activity of the NCRs in nodules (Haag et al., 2011). Besides their membrane-damaging activity, NCRs also have intracellular targets in rhizobia such as the cell division protein FtsZ and major chaperones like GroEL, suggesting that NCRs may interact with the bacterial machineries to induce bacterial differentiation (Farkas et al., 2014; Alunni and Gourion, 2016). In line with this hypothesis, treatment of bacterial cultures with sub-lethal doses of synthetic NCR peptides results in a partial mimicry of bacteroid differentiation, with bacterial cells that elongate and display an increased DNA content (Van de Velde et al., 2010). Finally, NCR peptides seem to interfere with the cell cycle regulatory network, as along their differentiation process, bacteroids display a major drop in the expression of the CtrA master regulator of the cell cycle and of its cognate regulon. A similar feature is also observed when synchronized bacterial cultures are treated with NCR peptides (Penterman et al., 2014; Roux et al., 2014; Pini et al., 2015).

Even if terminal bacteroid differentiation is rather the exception than the rule in the legume-rhizobium symbioses, it is found in at least five different clades (Oono and Denison, 2010). In the Dalbergioid clade, which comprises peanut (*Arachis hypogaea*) and *Aeschynomene* spp. plants, bacteroids become either elongated (E-type bacteroids) or even spherical (S-type bacteroids) depending on the host species (Staphorst and Strijdom, 1972; Bonaldi et al., 2011). These E- and S-type bacteroids are terminally differentiated, similarly as the ones found in IRLC legume nodules. In addition, the recent identification of NCR-like peptides in *Aeschynomene* plants and of CAPE peptides in peanut suggests that different plant clades use similar strategies to enforce terminal bacteroid differentiation to their rhizobial partners (Czernic et al., 2015; Karmakar et al., 2019). As NCR peptides from *Medicago* and *Aeschynomene* seem to derive from distinct ancestral genes, the most probable scenario is an evolutionary convergence between IRLC and Dalbergioids involving in each case the recruitment of immune peptides into the symbiotic process (Czernic et al., 2015). Interestingly, NCR genes are only expressed in nodules in both plant clades and seem to have functions that are now distinct from immune peptides as their expression is not pathogen responsive in *Medicago* (Guefrachi et al., 2014; Czernic et al., 2015).

Thus, plants that trigger terminal bacteroid differentiation invest a large amount of energy in maintaining a massive number of NCR genes in their genome and in the massive expression of these genes which may represent around 5% of all nodule transcripts. Moreover, ancestral state reconstruction indicated that ancestral bacteroids were of the U-morphotype. Thus the terminal bacteroid differentiation process observed in different clades of legumes must have arisen several times independently, suggesting that plants may draw some benefit from the terminal differentiated state of the bacteroids (Oono and Denison, 2010). However, what benefit plants could obtain from this costly process remained unclear (Mergaert et al., 2006). Comparison of pea (*Pisum sativum*, E-type bacteroids) and common bean (U-type bacteroids) nodules infected by *Rhizobium leguminosarum* strain A34 showed that E-type bacteroids displayed a higher symbiotic efficiency over U-type bacteroids (Oono and Denison, 2010). Similarly, *Bradyrhizobium* sp. 32H1 performed better in peanut (S-type bacteroids) than in cowpea (*Vigna unguiculata*, U-type) nodules. These studies suggest thus that bacteroid differentiation provides a higher return on investment to the host plant (S > U, E > U). In line with this hypothesis, a recent study of the symbiotic efficiency of two *M. truncatula* accessions in interaction with four *Sinorhizobium* spp. strains equally suggested the existence of a positive correlation between the extent of bacteroid differentiation (i.e., the ploidy level of the bacteroids and their cell size) and the symbiotic efficiency of the interaction (Kazmierczak et al., 2017). Using the *Aeschynomene-Bradyrhizobium* system, we recently demonstrated a higher symbiotic efficiency of *Bradyrhizobium* strain ORS285 as S-type bacteroids in *Aeschynomene indica* nodules than as E-type bacteroids in *Aeschynomene afraspera* nodules (S > E) by combining biomass measurements and C/N elemental and fluxomic analyses (Lamouche et al., 2018). In this experimental setup, S-type bacteroids display a higher ploidy than the E-type ones, suggesting that S-type bacteroids are more differentiated than E-type ones. Moreover, the

formation of S-type bacteroids in *A. indica* involves an intermediate step of cell elongation before cell polarity is lost, demonstrating that the S-type bacteroids are a further developmental stage (Czernic et al., 2015).

To broaden the scope of this latter comparison, we wanted to know if the higher efficiency of S-type over E-type bacteroids and if the correlation between symbiotic efficiency and the extent of bacteroid differentiation are valid in a larger sample of Dalbergioid-*Bradyrhizobium* associations (**Figure 1A**). We used a combination of *Bradyrhizobium* spp. that can nodulate an E-type bacteroid-forming plant (*A. afraspera*) and/or S-type bacteroid-inducing plants (*A. indica* and *Aeschynomene evenia*) (Molouba et al., 1999). We show that the more strongly differentiated S-type bacteroids are more efficient than the E-type ones. Nevertheless, at a finer scale, among the S-type bacteroids only, no correlation exists anymore between symbiotic efficiency and the level of differentiation (bacteroid ploidy level and cell size), suggesting the existence of other factors that contribute to efficiency.

MATERIALS AND METHODS

Plant Growth and Inoculation

Bradyrhizobium sp. ORS278, ORS285, ORS287, ORS335, ORS357, and BTAi1 were grown at 30°C in YM medium. *A. afraspera*, *A. indica (Africa n°19)*, and *A. evenia* ssp. *evenia (n°21)* seeds were surface-sterilized by soaking the seeds for 45 min in 100% sulfuric acid for scarification. After 5 washes in sterile H_2O, seed coat sterilization was performed for 30 min in 0.38% bleach followed by 10 washes in sterile H_2O. Seeds were germinated in the dark on 0.7% (m/v) Kalys agar plates at 30°C (Guefrachi et al., 2015). Plantlets were transferred and grown in a semi-sterile system composed of test tubes filled

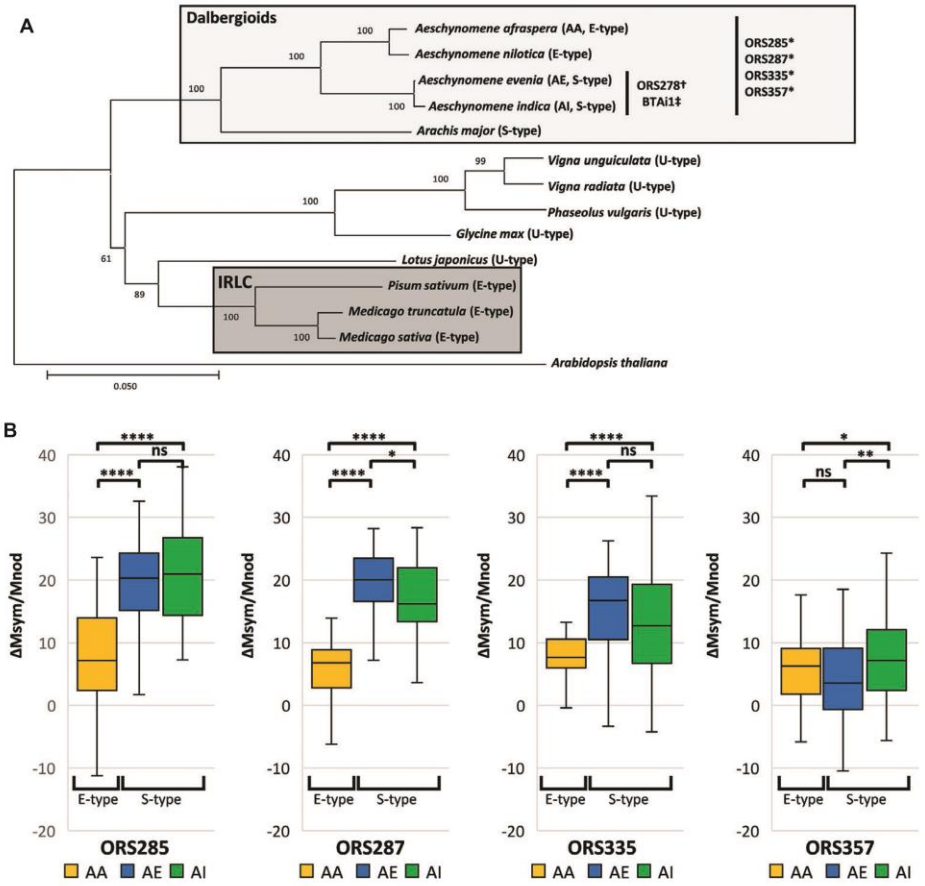


FIGURE 1 | S-type bacteroids display higher symbiotic efficiency than E-type bacteroids in *Bradyrhizobium-Aeschynomene* symbioses. **(A)** Neighbor-joining tree based on nucleic *matK* sequences from Azani *et al.* (2017) of selected legume species displaying contrasting bacteroid morphotypes. The distant IRLC and Dalbergioid clade both harbor differentiated bacteroids (E- and E-/S-type, respectively) whereas other groups harbor unmodified bacteroids (U-type). Compatible associations between *Aeschynomene* spp. and *Bradyrhizobium* strains used in this study are mentioned. Strains were isolated from *: *A. afraspera*; †: *A. sensitiva*, ‡: *A. indica* (Molouba et al., 1999). **(B)** Mass gains per gram of dry nodule at 14 dpi of inoculated plants over non-inoculated (n ≥ 30) respectively elicited by *Bradyrhizobium* sp. ORS285, ORS287, ORS335, and ORS357. Significant differences after ANOVA and *post hoc* Tukey tests are indicated as ns: $p > 0.05$; *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$; ****: $p \leq 0.0001$. AA: *A. afraspera*; AE: *A. evenia*; AI: *A. indica*.

with BNM medium and covered with an aluminum foil displaying a small hole for the tap root to grow through it (Bonaldi et al., 2011). Thus, the shoot is growing outside the tube whereas the root system is bathing in the culture medium. Three days after transfer, plants were inoculated with 2 ml of bacterial suspension (10^8 cfu/ml) of the corresponding strains and cultivated during 14 days (Lamouche et al., 2018). Plants were grown at 28°C with 80% humidity under a 16 h:8 h of light:dark regimen. For 45 days post-inoculation (dpi) observations, plants were transferred in 2-L bottles and grown in the same climatic chambers.

Microscopic Observation of Nodule Sections

Fourteen dpi nodules were harvested and 70 µm sections were produced using a Leica VT1200S vibratome before staining for 20 min with the LIVE/DEAD BacLight Bacterial viability kit (0.5 µl PI, 0.5 µl SYTO9, Thermo) in 50 mM Tris-HCl pH 7.0 buffer containing 0.01% calcofluor white M2R (Sigma). Nodule sections were rinsed and mounted in 50 mM Tris-HCl pH 7.0 buffer and imaged on a Leica SP8 confocal microscope.

Plant Biomass Estimations

Plants ($n \geq 30$ per condition) were collected at 14 dpi. Nodules, roots, and shoots were separated and dried at 80°C for 48 h. Dry mass was measured for the shoot (M_{shoot}), root (M_{root}), and nodule (M_{nod}) plant compartments, the sum of M_{shoot} and M_{root} being called M_{inoc} for inoculated plants and M_{mock} for non-inoculated control plants. The proxy used to measure symbiotic efficiency was the total plant dry mass gain per mg of nodule, calculated as the dry mass difference between the inoculated and non-inoculated plants ($\Delta M_{sym} = M_{inoc} - M_{mock}$), divided by the nodule dry mass (i.e., Efficiency = $\Delta M_{sym}/M_{nod}$). Significant differences were determined by analysis of variance followed by a *post hoc* HSD Tukey test ($p < 0.05$).

Determination of Nitrogen Fixation Activity by Acetylene Reduction Assays

Fourteen dpi nodulated root systems were harvested and inserted into rubber-sealed glass vials containing 250 µl of sterile water. Vials were incubated for 2 h with 1 ml of acetylene at room temperature in the dark and 250 µl of gas was removed and injected in a gas chromatography system (7820A, Agilent). The area under the curve was measured for the ethylene peak, and integrated in the calculation of nitrogenase activity as previously described (Barrière et al., 2017). Dry mass measurements were performed in parallel on the same samples. More than 30 plants were used per condition.

Flow Cytometry Analyses

For each condition, nodules were harvested on more than 15 inoculated plants and pooled to have a representative sample for flow cytometry analyses. Nodules were crushed with a mortar and a pestle in a bacteroid extraction buffer (BEB: 125 mM KCl, 50 mM Na-succinate, 50 mM TES, 0.1% BSA, pH 7.0). The nodule homogenate was centrifuged for 10 min at 100*g* to pellet plant debris. The supernatant containing the bacteroids was centrifuged for 10 min at 3600*g* to pellet the bacteroids before resupsension in a small volume of BEB. All the procedures were performed at 4°C. Bacteroids were heat-killed (for 10 min at 70°C) and stained using 50 µg ml^{-1} propidium iodide (PI). Cell size (forward scatter of the laser ray) and DNA content (level of fluorescence of the PI) were determined on a MoFlo Astrios flow cytometer and the results were analyzed using the Summit 6.2 software (Beckman Coulter).

RESULTS AND DISCUSSION

Bradyrhizobium-Aeschynomene Associations Lead to the Formation of Functional Nodules Containing E-Type and S-Type Bacteroids Depending on the Host Plant

Before analyzing the symbiotic efficiency of E-type and S-type bacteroids, we first determined the functionality of the symbiosis in the 16 associations formed by *A. afraspera*, *A. evenia*, and *A. indica* and 6 *Bradyrhizobium* strains (BTAi1, ORS278, ORS285, ORS287, ORS335, and ORS357; the two first strains are unable to form nodules on *A. afraspera* because they lack the required Nod factors for nodulation of these plants (Giraud et al., 2007; Chaintreuil et al., 2013). In all cases, functional nodules were formed which displayed a pink coloration of the infected zone due to leghemoglobin production as observed in the hand-sectioned nodules, although the coloration was paler in the case of *A. evenia* nodulated by *Bradyrhizobium* sp. ORS357 (**Figure S1**). In agreement with this observation, confocal imaging of nodule sections showed for each interaction the presence of a large central zone composed of plant symbiotic cells infected by rhizobia (**Figure S2**). Bacteroid differentiation into S- or E-type bacteroids was confirmed in all plant species/bacterial strain combinations by confocal imaging at high magnification. In agreement with the bacteroid type being determined by the host, bacteroids were spherical in *A. evenia* and *A. indica*, and elongated in *A. afraspera* regardless of the bacterial strain that was inoculated (**Figure 2**). Thus, all the interactions tested in this study lead to the formation of functional nodules where rhizobia undergo terminal differentiation into S- or E-type bacteroids depending on the host plant. However, this gross estimation of symbiotic functionality is only qualitative and does not discriminate between weakly and highly efficient associations.

Spherical Bacteroids Display Higher Symbiotic Efficiency Than Elongated Bacteroids in the *Bradyrhizobium/Aeschynomene* Symbiosis

We estimated the symbiotic efficiency for the four "generalist" strains (ORS285, ORS287, ORS335, ORS357) that can interact with both E-type bacteroid-inducing plants (*A. afraspera*) and S-type bacteroid-inducing plants (*A. evenia* and *A. indica*). To do so, we determined the gain in plant biomass attributable

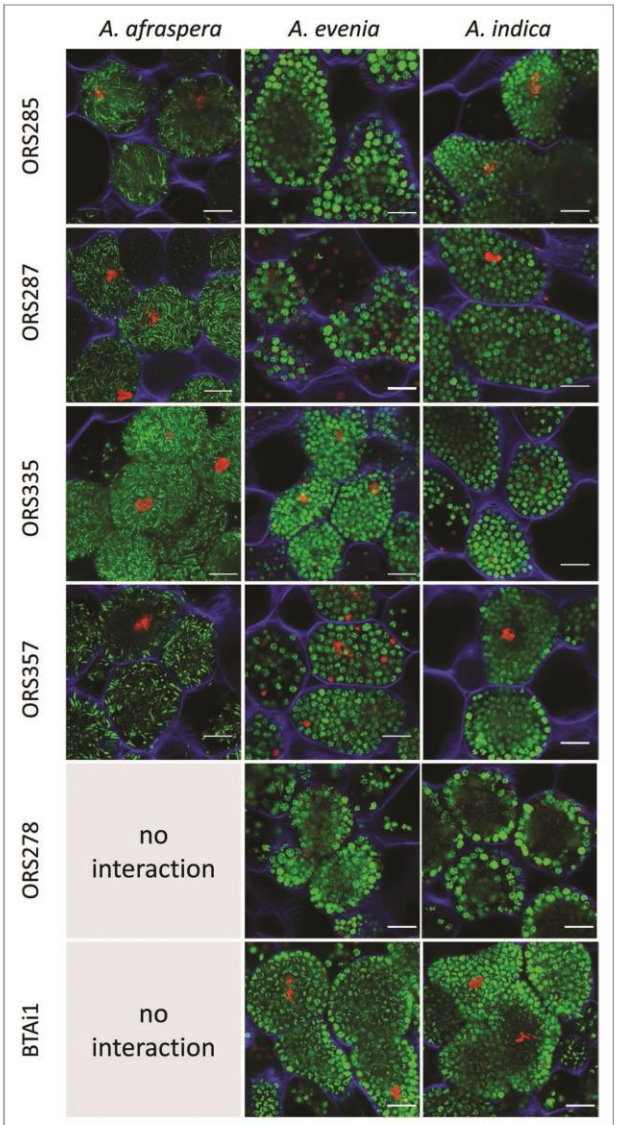


FIGURE 2 | *A. indica* and *A. evenia* induce S-type bacteroids whereas *A. afraspera* induces E-type bacteroids in symbiosis with all the tested strains. Fourteen dpi nodule sections stained with the LIVE/DEAD kit (live bacteria appear green, dead bacteria and plant nuclei appear red) and calcofluor white (plant cell walls appear blue) display differentiated bacteroids within plant cells. In *A. afraspera*, bacteroids adopt an elongated morphotype, whereas in *A. evenia* and *A. indica*, bacteroids adopt a spherical morphotype regardless of the bacterial strain inoculated on the plants. Scale bars: 10 µm.

to the symbiosis and per investment in the symbiosis. The dry biomass of the different plant parts (i.e., shoots, roots, and nodules) was measured for both inoculated and non-inoculated plants for the different symbiotic couples at 14 dpi. The total dry mass of non-inoculated plants (M_{mock}) was subtracted from the mass of the inoculated plants (M_{inoc}). The resulting value (ΔM_{sym}) thus corresponds to the biomass gained through symbiosis because it corrects the biomass of the inoculated plants by the portion of the biomass derived from the usage of nutrients stored in seeds that are different in size between *A. afraspera* (producing large seeds) and *A. indica* and *A. evenia* (producing smaller seeds). Moreover, since the onset of nitrogen fixation occurs at the same time (at 5 dpi) in the different *Aeschynomene* species, independently of the seed size and plantlet sizes (Bonaldi et al., 2011), all the tested plants fixed nitrogen during a same period (9 days) at the time of harvesting (14 dpi). Finally, the symbiotic efficiency was obtained by dividing ΔM_{sym} with the nodule dry mass (M_{nod}) which is an estimation of the investment of the plant in the symbiosis. We previously used this parameter (i.e., $\Delta M_{sym}/M_{nod}$) to compare *A. afraspera* and *A. indica* plants in symbiosis with *Bradyrhizobium* strain ORS285 (Lamouche et al., 2018).

The use of the plant biomass parameter for estimating the efficiency of the symbiotic process has the advantage to be a time-integrated measurement and thus to be less sensitive to temporary fluctuations that might affect instant measurements of the nitrogenase activity, for example through the often used acetylene reduction assays (ARA). We studied the relation between plant biomass parameters and the activity of nodules in the ARA using the strain ORS285 on the three studied host plants. We found that the total plant dry weight was correlated to ARA measurements in the three conditions and that $\Delta M_{sym}/M_{nod}$ was correlated to the ARA measurements in two of the three hosts, with *A. evenia* showing more variability in the mass measurements (**Figure S3**). Thus, we considered that $\Delta M_{sym}/M_{nod}$ is a good proxy for assessing the efficiency of the symbiotic process.

We found that *A. indica* and *A. evenia* plants, inducing S-type bacteroids, obtained a significantly higher benefit ($p < 0.0001$) than the *A. afraspera* plants inducing E-type bacteroids during symbiosis with all tested strains except strain ORS357 (**Figure 1B**). However, the latter strain did not seem to be efficiently interacting with plants inducing S-type bacteroids as suspected from the pale color of the nodule inner tissues (**Figure S1**). Thus overall, S-type bacteroids seem to be more efficient than the E-type ones in the Dalbergioid clade, extending our previous observations that were based on the performance of strain ORS285 in *A. afraspera* and *A. indica* (Lamouche et al., 2018). The high efficiency of the S-type bacteroids is further confirmed by the biomass of plants grown until 45 dpi when growth of *A. evenia* and *A. indica* plants had exceeded growth of *A. afraspera* plants, despite the fact that the former two species form initially smaller plantlets from smaller seeds than the latter one (**Figure S4**).

Comparison of the Symbiotic Efficiency of Different Bacterial Strains on a Single Host Plant

We then considered the efficiency of a single plant genotype interacting with different symbiont strains. In all 16 interactions tested, the shoot biomass of inoculated plants was significantly higher than the non-inoculated controls, confirming that the symbiosis is a major driver of plant growth in case of nitrogen starvation (**Table 1**). This general gain in biomass attributable

TABLE 1 | Strain-specific assessment of symbiotic efficiency in *Aeschynomene* spp.

	Shoot mass (M_{shoot}) (mg)	Root mass (M_{root}) (mg)	Nodule mass (M_{nod}) (mg)	Nodule number	Total mass (M_{tot}) (mg)	Shoot/root mass ratio	Efficiency ($\Delta M_{sym}/M_{nod}$)
A. afraspera							
ORS285	63 (±2.38)^{a}	20 (±0.9)^{a}	2.6 (±0.12)^{c}	24 (±1.22)^{a}	86 (±3.23)^{ab}	3.3 (±0.13)^{a}	7.82 (±1.46)^{b}
ORS287	64 (±2.12)^{a}	20 (±0.88)^{a}	3.14 (±0.13)^{a}	20.8 (±0.84)^{a}	86 (±2.97)^{ab}	3.35 (±0.11)^{a}	2.8 (±1.03)^{a}
ORS335	71 (±2.2)^{a}	22 (±0.83)^{a}	3.65 (±0.14)^{b}	23.03 (±0.96)^{a}	97 (±3.03)^{a}	3.25 (±0.06)^{ab}	5.53 (±0.83)^{ab}
ORS357	63 (±2.74)^{a}	22 (±0.81)^{a}	2.9 (±0.13)^{ac}	20.63 (±1.17)^{a}	87 (±3.55)^{ab}	2.91 (±0.06)^{b}	3.33 (±1)^{a}
BTAi1	—	—	—	—	—	—	—
ORS278	—	—	—	—	—	—	—
NI	46 (±2.71)^{b}	28 (±1.7)^{b}	—	—	74 (±4.29)^{b}	1.74 (±0.08)^{c}	—
A. evenia							
ORS285	28 (±0.89)^{ab}	8 (±0.22)^{a}	0.77 (±0.03)^{ab}	13.95 (±0.48)^{ab}	37 (±1.08)^{ab}	3.68 (±0.1)^{a}	18.95 (±0.97)^{a}
ORS287	30 (±0.76)^{ac}	8 (±0.21)^{a}	0.86 (±0.03)^{ab}	15.45 (±0.43)^{a}	39 (±0.96)^{a}	3.86 (±0.06)^{a}	20.31 (±1.06)^{a}
ORS335	27 (±0.8)^{b}	7 (±0.23)^{ab}	0.8 (±0.03)^{ab}	14.18 (±0.51)^{ab}	34 (±0.98)^{b}	3.89 (±0.16)^{a}	15.16 (±0.95)^{a}
ORS357	17 (±0.44)^{d}	6 (±0.18)^{b}	0.67 (±0.02)^{b}	16.15 (±0.44)^{a}	24 (±0.61)^{c}	2.75 (±0.04)^{b}	3.66 (±0.9)^{b}
BTAi1	26 (±0.84)^{b}	12 (±0.34)^{c}	0.67 (±0.05)^{b}	11.45 (±0.94)^{b}	39 (±1.13)^{a}	2.17 (±0.06)^{c}	30.28 (±2.58)^{c}
ORS278	33 (±0.93)^{c}	10 (±0.26)^{e}	1.39 (±0.04)^{c}	34.5 (±1.11)^{c}	44 (±1.18)^{d}	3.33 (±0.06)^{e}	16.09 (±0.52)^{a}
NI	12 (±0.43)^{e}	9 (±0.29)^{d}	—	—	21 (±0.64)^{c}	1.43 (±0.04)^{d}	—
A. indica							
ORS285	47 (±1.62)^{a}	14 (±0.42)^{ac}	1.29 (±0.05)^{a}	18.12 (±0.59)^{a}	62 (±2)^{a}	3.45 (±0.07)^{a}	20.33 (±1.08)^{a}
ORS287	48 (±1.57)^{ab}	13 (±0.47)^{abc}	1.62 (±0.05)^{b}	21.32 (±0.79)^{b}	63 (±2.01)^{a}	3.71 (±0.07)^{a}	16.62 (±0.84)^{ab}
ORS335	39 (±1.33)^{c}	11 (±0.37)^{c}	1.27 (±0.04)^{a}	17.78 (±0.56)^{a}	52 (±1.69)^{b}	3.61 (±0.08)^{a}	12.77 (±1.02)^{bc}
ORS357	33 (±1.16)^{d}	12 (±0.52)^{ac}	1.41 (±0.06)^{ab}	21.33 (±0.72)^{b}	46 (±1.59)^{b}	2.8 (±0.06)^{b}	7.32 (±0.9)^{d}
BTAi1	32 (±1.15)^{d}	14 (±0.51)^{b}	1.27 (±0.05)^{a}	14.83 (±0.59)^{c}	48 (±1.53)^{b}	2.28 (±0.06)^{c}	9.39 (±1.02)^{cd}
ORS278	53 (±1.81)^{b}	17 (±0.63)^{d}	1.88 (±0.06)^{c}	27.87 (±0.79)^{d}	72 (±2.4)^{d}	3.08 (±0.07)^{b}	19.64 (±1.08)^{a}
NI	17 (±1.02)^{e}	17 (±0.64)^{d}	—	—	34 (±1.35)^{c}	1.07 (±0.07)^{d}	—

Shoot mass, root mass, nodule mass, shoot/root ratio, and symbiotic efficiency of the different Bradyrhizobium-Aeschynomene interactions measured at 14 dpi. NI stands for non-inoculated control plants. Mean values are displayed and the standard error is indicated between brackets. Letters represent significant differences after ANOVA and post hoc Tukey HSD tests ($p < 0.05$) performed independently on each plant.

to the symbiosis indicates that the interaction is functional and fuels plant growth with nitrogen. However, the root mass was similar or even higher in non-inoculated plants than in the nodulated plants, suggesting that non-inoculated plants invest more in root growth. This is a well-known response of nitrogen-starved plants that increase their root growth to explore additional areas in the soil for nitrogen sources, resulting in a trade-off between shoot growth and root growth (Wilson, 1988). Accordingly, the shoot/root (S/R) ratio, which is often used as a proxy for the nutritional status of the plants, was higher for inoculated plants compared to control plants (**Table 1**). For *A. afraspera*, the differences between strains was limited, with shoot and root biomass giving similar results although the ORS285 gave the highest efficiency and ORS357 the lowest S/R ratio (**Table 1**).

More contrasted results were observed in S-type bacteroid-forming plants. While ORS278-infected plants displayed the highest total masses, they also have high nodule masses lowering the global symbiotic efficiency (**Table 1**). BTAi1 had the highest calculated efficiency on *A. evenia* but the S/R ratio remained low, suggesting that the plant's nitrogen needs were not fulfilled. In this particular case, the apparent high efficiency of BTAi1 could be partially distorted by the high root mass. In addition, this contradicting observation could be also due in part to the low number of nodules formed by this strain which were however bigger than the nodules formed by the other strains (+45% in individual nodule mass as compared to ORS278).

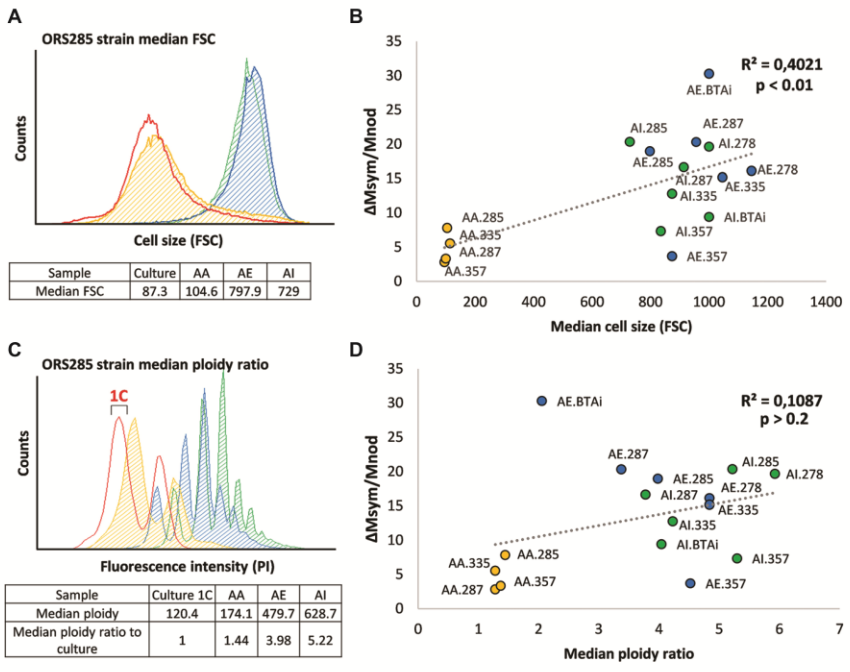
Regarding the symbiotic efficiency on the three *Aeschynomene* species used in this study, ORS285 appeared to be the most versatile strain, with high efficiency on all tested plants. On the other hand, ORS357 seemed to be quite efficient in symbiosis with *A. afraspera,* which gives rise to E-type bacteroids, but not with the two plants harboring S-type bacteroids. Strains ORS287 and ORS335 showed intermediate values of symbiotic efficiency with the three host plants. Additionally, considering the original plant from which the bacterial strains were isolated (i.e., BTAi1 and ORS278 from S-type *Aeschynomene* and ORS285, ORS287, ORS335, and ORS357 from E-type *Aeschynomene*), we cannot link symbiotic efficiency to a co-evolution/host adaptation process that would improve the output of the symbiosis for the plant. As in the IRLC, the extent of the NCR repertoire in the plant may affect the quality of the interaction with a given strain, and on the bacterial side, the capacity to cope with NCR peptides may differ between strains. Moreover, it was shown in *M. truncatula* that a single plant genotype responds to different compatible rhizobial strains by expressing NCR genes differentially (Nallu et al., 2014; Kazmierczak et al., 2017). Similarly, *A. afraspera* expressed NCR genes to different levels when plants were inoculated by *Bradyrhizobium diazoefficiens* USDA110 or *Bradyrhizobium* sp. ORS285 (Barrière et al., 2017), suggesting that differential expression of NCR genes may take place in the different plant species/bacterial strain combinations tested here. Alternatively, the differences observed here may be due to the membrane properties of the different

bacterial strains, and/or to a differential affinity of the BclA transporters present in the bacterial strains to the cocktails of NCR peptides produced by the different *Aeschynomene* plants used in this study.

The Bacteroid Differentiation Level Is Not the Only Determinant of Symbiotic Efficiency in *Aeschynomene-Bradyrhizobium* Associations

The differentiation level of the bacteroids was determined by calculating both their median ploidy ratio when compared to a bacterial culture used as reference for the haploid state (1C) (**Figure 3A**), as well as their cell size determined by the forward light scatter (**Figure 3C**). For each of the four generalist strains, we observed that both S-type bacteroids displayed higher ploidy levels, cell size, and symbiotic efficiency than E-type ones, in agreement with the previously published data on ORS285 (**Figures 3B,D**; Czernic et al., 2015; Lamouche et al., 2018). Taking together the results of all strains revealed a differentiation gradient from *A. afraspera*, *A. evenia* to *A. indica* bacteroids, except for ORS335 bacteroids that displayed a higher ploidy level in *A. evenia* nodules (**Figure 3B**). We then searched for a potential correlation between the bacteroid differentiation level and the symbiotic efficiency as observed in the IRLC (Kazmierczak et al., 2017). While increased bacteroid ploidy and cell size are correlated in our analyzed *Aeschynomene-Bradyrhizobium* associations (R^2 = 0.63, p = 2.5e^{-4}), the correlation between symbiotic efficiency and these two hallmarks of differentiation is observed only for cell size (**Figure 3B**; R^2 = 0.40; p = 4.3e^{-3}) and not for ploidy (**Figure 3D**; R^2 = 0.11; p = 0.21). However, the latter correlation analysis is influenced by some conditions that have an unexpected and opposite relationship between efficiency and differentiation. Strain BTAi1 symbiosis with *A. evenia* was highly efficient according to our estimation while the bacteroids displayed a low endoreduplication rate, although as argued above, the efficiency estimation for strain BTAi1 in *A. evenia* might be biased. On the contrary, strain ORS357, which exhibited a low symbiotic efficiency in S-type bacteroids, underwent a high degree of differentiation.

Taken together, the correlation between bacteroid differentiation level and the symbiotic efficiency might be affected by other factors that contribute to the optimal functioning of the bacteroids. In the IRLC, it was shown that *Sinorhizobium fredii* HH103 does not undergo terminal differentiation in the NCR-producing *Glycyrrhiza uralensis*



Sample	Culture	AA	AE	AI
Median FSC	87.3	104.6	797.9	729

Sample	Culture 1C	AA	AE	AI
Median ploidy	120.4	174.1	479.7	628.7
Median ploidy ratio to culture	1	1.44	3.98	5.22

FIGURE 3 | Relationship between symbiotic efficiency and bacteroid differentiation in *Bradyrhizobium-Aeschynomene* symbioses. **(A,C)** Examples of flow cytometry analysis of *Bradyrhizobium* sp. Strain ORS285 using, respectively, the forward scatter as an estimation of cell size **(A)** and propidium iodide fluorescence to estimate bacteroid ploidy **(C)** for the free-living culture reference (red), and for the *A. afraspera* (orange), *A. evenia* (blue), and *A. indica* (green) bacteroids. The fluorescence value of the 1C peak (see bracket) representing one haploid genome is the reference used to measure the median ploidy ratio of bacteroids. E-type bacteroids display a very moderate differentiation while S-type bacteroids are very different from the culture. **(B,D)** Correlation estimations between symbiotic efficiency (measured as the mass gain per gram of dry nodule) and the median cell size **(B)** or median ploidy ratios observed **(D)**. A significant correlation is only observed between bacteroid cell size and symbiotic efficiency. AA: *A. afraspera*; AE: *A. evenia*; AI: *A. indica*.

nodules although it fixes efficiently nitrogen and supports plant growth (Crespo-Rivas et al., 2016; Montiel et al., 2017). The bacterium is able to modify its LPS *in planta* possibly to cope with NCR-induced membrane stress. Similarly, some strains in our panel may somehow resist the differentiation program and investigating the structure of their LPS or other membrane components like the hopanoids (Silipo et al., 2014; Kulkarni et al., 2015) may explain partly why low differentiation levels can be associated to a high efficiency. It is also possible that the genetic differences between rhizobial strains result in the production of signals that induce strain-specific responses from the plant. As mentioned above, nodules express NCR genes to different relative levels in response to infections with different rhizobium strains. Although we did not observe any obvious signs of defense reactions in the nodules, it cannot be excluded that plants may induce an immune response to various extents against some strains, resulting in a reduced symbiotic metabolism regardless of the bacteroid differentiation level that is reached. Another aspect is that NCR peptides have been shown to have multiple intracellular targets in bacteroids in the *Medicago-Sinorhizobium* symbiosis (Farkas et al., 2014). We can speculate that similarly, *Aeschynomene* NCR peptides interact with different sets of targets in the different *Bradyrhizobium* strains affecting, thereby, the functioning of the bacteroids. Finally, as NCR peptides have a direct action on both the bacterial envelope and intracellular targets that may affect the cell metabolism, it is possible that the cocktail of NCR peptides produced by *Aeschynomene* plants inducing S-type bacteroids has an opposing effect on differentiation and symbiotic efficiency in some strains (Alunni and Gourion, 2016).

CONCLUSION

In IRLC plants, it was observed that the level of bacteroid differentiation is correlated with the size of the cationic NCR repertoire (Montiel et al., 2017). In Dalbergioids, limited genomic data are available but it is not unlikely that the NCR repertoire is both quantitatively and qualitatively different between *Aeschynomene* plants inducing E- or S-type bacteroids (Czernic et al., 2015). Hence, it is possible that the NCR cocktails secreted by S-type bacteroid-forming plants exert a more stringent action on the bacteroid cell cycle and metabolism than E-type bacteroid-forming *Aeschynomene*, constraining the symbionts to a higher productivity beneficial to the host. On the other hand, while our results show that NCR-mediated differentiation affects the efficiency of the symbiotic process, additional yet unknown factors are modulating symbiotic efficiency in *Aeschynomene* spp.

AUTHOR CONTRIBUTIONS

FL, PM, and BA conceived the experiment(s), analyzed the result(s), and wrote the paper. FL, NB-B, and BA conducted the experiment(s). All authors reviewed the manuscript.

FUNDING

This research was funded by the Agence Nationale de la Recherche (ANR-17-CE20-0011-02) and the Labex "Saclay Plant Sciences" (ANR-10-LABX-0040-SPS). FL was supported by a PhD fellowship of the Université Paris-Sud (L866-868).

ACKNOWLEDGMENTS

We thank Dr. Eric Giraud and Joël Fardoux for providing *Bradyrhizobium* sp. ORS287, ORS335, ORS357 strains and *Aeschynomene* seeds, as well as Corine Foucault and Stacy Simoes for technical help.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00377/full#supplementary-material

FIGURE S1 | *Aeschynomene-Bradyrhizobium* associations form functional nodules. Hand-made sections of nodules show the inner part of the symbiotic organ that display a red color indicating that leghemoglobin, a marker of nodule functioning, is produced. Scale bars: 1 mm.

FIGURE S2 | Functional symbiosis is associated with the presence of central nodule tissues composed of symbiotic cells infected by rhizobia. Fourteen dpi nodules were sectioned (70 µm) and stained with LIVE/DEAD kit (live/dead bacteria appear green/red respectively) and calcofluor white (plant cell walls appear blue) before confocal imaging. Scale bars: 100 µm.

FIGURE S3 | Plant biomass parameters correlate with nitrogen fixation activity in symbiotic systems involving *Bradyrhizobium* sp. ORS285 and *A. afraspera or A. indica* and to a lesser extent with *A. evenia*. Relationship between plant total dry mass (left panel) or symbiotic efficiency ($\Delta M_{sym}/M_{nod}$, right panel) and nitrogenase activity (analyzed by acetylene reduction assay and expressed as nmol C_2H_2 evolved per hour and per plant) was tested by calculating the correlation coefficient (R^2) and its associated p-value for 14 dpi nodulated plants of *A. afraspera* **(A)**, *A. evenia* **(B)** and *A. indica* **(C)**.

FIGURE S4 | General aspect of *Aeschynomene* plants inducing E- and S-type bacteroids at 45 dpi with ORS285 **(A)**, ORS335 **(B)**, ORS287 **(C)** or ORS357 **(D)**. The small seed producing *A. evenia* and *A. indica* developed a higher shoot length than the large seed producing *A. afraspera* regardless of the symbiotic strain. Scare bars: 5 cm.

REFERENCES

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. doi: 10.1111/nph.14025

Alunni, B., Kevei, Z., Redondo-Nieto, M., Kondorosi, A., Mergaert, P., and 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. doi: 10.1094/MPMI-20-9-1138

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. doi: 10.1038/s41598-017-08830-0

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. doi: 10.1094/MPMI-04-11-0093

Chaintreuil, C., Arrighi, J. F., Giraud, E., Miché, L., Moulin, L., Dreyfus, B., et al. (2013). Evolution of symbiosis in the legume genus *Aeschynomene*. *New Phytol.* 200, 1247–1259. doi: 10.1111/nph.12424

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. doi: 10.1111/1462-2920.13101

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., et al. (2015). Convergent evolution of rndosymbiont differentiation in dalbergioid and inverted repeat-lacking clade legumes mediated by nodule-specific cysteine-rich peptides. *Plant Physiol.* 169, 1254–1265. doi: 10.1104/pp.15.00584

Farkas, A., Maroti, G., Durg, H., Gyorgypal, Z., Lima, R. M., Medzihradszky, K. F., et al. (2014). *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* 111, 5183–5188. doi: 10.1073/pnas.1404169111

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.-C., et al. (2007). Legumes symbioses: absence of nod genes in photosynthetic bradyrhizobia. *Science* 316, 1307–1312. doi: 10.1126/science.1139548

Gruber, N., and Galloway, J. N. (2008). An Earth-system perspective of the global nitrogen cycle. *Nature* 451, 293–296. doi: 10.1038/nature06592

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 Genom.* 15:712. doi: 10.1186/1471-2164-15-712

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. doi: 10.1094/MPMI-04-15-0094-R

Haag, A. F., Baloban, M., Sani, M., Kerscher, B., Pierre, O., Farkas, A., et al. (2011). Protection of *Sinorhizobium* against host cysteine-rich antimicrobial peptides is critical for symbiosis. *PLoS Biol.* 9:e1001169. doi: 10.1371/journal.pbio.1001169

Horváth, B., Domonkos, Á., Kereszt, A., Szűcs, A., Ábrahám, E., Ayaydin, F., et al. (2015). Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the *Medicago truncatula dnf7* mutant. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15232–15237. doi: 10.1073/pnas.1500777112

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, 271–285. doi: 10.1094/MPMI-06-18-0174-R

Kazmierczak, T., Nagymihály, M., Lamouche, F., Barrière, 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. doi: 10.1094/MPMI-01-17-0009-R

Kim, M., Chen, Y., Xi, J., Waters, C., Chen, R., and Wang, D. (2015). An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15238–15243. doi: 10.1073/pnas.1500123112

Kondorosi, E., Mergaert, P., and Kereszt, A. (2013). A paradigm for endosymbiotic life: cell differentiation of *Rhizobium* bacteria provoked by host plant factors. *Annu. Rev. Microbiol.* 67, 611–628. doi: 10.1146/annurev-micro-092412-155630

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. doi: 10.1128/mBio.01251-15

Lamouche, F., Gully, D., Chaumeret, A., Nouwen, N., Verly, C., Pierre, O., et al. (2018). Transcriptomic dissection of *Bradyrhizobium* sp. strain ORS285 in symbiosis with *Aeschynomene* spp. inducing different bacteroid morphotypes with contrasted symbiotic efficiency. *Environ. Microbiol.* doi: 10.1111/1462-2920.14292

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, 161–173. doi: 10.1104/pp.102.018192

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. doi: 10.1073/pnas.0600912103

Mikuláss, K. R., Nagy, K., Bogos, B., Szegletes, Z., Kovács, E., Farkas, A., et al. (2016). Antimicrobial nodule-specific cysteine-rich peptides disturb the integrity of bacterial outer and inner membranes and cause loss of membrane potential. *Ann. Clin. Microbiol. Antimicrob.* 15, 1–5. doi: 10.1186/s12941-016-0159-8

Molouba, F., Lorquin, J., Willems, A., Hoste, B., Giraud, E., Dreyfus, B., et al. (1999). Photosynthetic bradyrhizobia from *Aeschynomene* spp. are specific to stem-nodulated species and form a separate 16S ribosomal DNA restriction fragment length polymorphism group. *Appl. Environ. Microbiol.* 65, 3084–3094.

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. doi: 10.1073/pnas.1704217114

Nallu, S., Silverstein, K. A. T., Zhou, P., Young, N. D., and Vandenbosch, K. A. (2014). Patterns of divergence of a large family of nodule cysteine-rich peptides in accessions of *Medicago truncatula*. *Plant J.* 78, 697–705. doi: 10.1111/tpj.12506

Oono, R., and Denison, R. F. (2010). Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol.* 154, 1541–1548. doi: 10.1104/pp.110.163436

Pecrix, Y., Staton, S. E., Sallet, E., Lelandais-Brière, C., Moreau, S., Carrère, S., et al. (2018). Whole-genome landscape of *Medicago truncatula* symbiotic genes. *Nat. Plants* 4, 1017–1025. doi: 10.1038/s41477-018-0286-7

Penterman, J., Abo, R. P., De Nisco, N. J., Arnold, M. F. F., Longhi, R., Zanda, M., et al. (2014). Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3561–3566. doi: 10.1073/pnas.1400450111

Pini, F., De Nisco, N. J., Ferri, L., Penterman, J., Fioravanti, A., Brilli, M., et al. (2015). Cell cycle control by the master regulator CtrA in *Sinorhizobium meliloti*. *PLos Genet.* 11:e1005232. doi: 10.1371/journal.pgen.1005232

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, 817–837. doi: 10.1111/tpj.12442

Silipo, A., Vitiello, G., Gully, D., Sturiale, L., Chaintreuil, C., Fardoux, J., et al. (2014). Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. *Nat. Commun.* 5:5106. doi: 10.1038/ncomms6106

Sprent, J. I., Ardley, J., and James, E. K. (2017). Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytol.* 215, 40–56. doi: 10.1111/nph.14474

Staphorst, J. L., and Strijdom, B. W. (1972). Some observations on the bacteroids in nodules of *Arachis* sp. and the isolation of *Rhizobia* from these nodules. *Phytophylactica* 4, 87–92.

The Legume Phylogeny Working Group (LPWG) (2017). A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. *Taxon* 66, 44–77. doi: 10.12705/661.3

Tiricz, H., Szücs, 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. doi: 10.1128/AEM.01791-13

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. doi: 10.1126/science.1184057

Wang, Q., Yang, S., Liu, J., Terecskei, K., Ábrahám, E., Gombár, A., et al. (2017). Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. *Proc. Natl. Acad. Sci. U.S.A.* 114, 6854–6859. doi: 10.1073/pnas.1700715114

Wilson, J. B. (1988). A review of evidence on the control of shoot: root ratio, in relation to models. *Ann. Bot.* 61, 433–449. doi: 10.1093/oxfordjournals.aob.a087575

Yang, S., Wang, Q., Fedorova, E., Liu, J., Qin, Q., Zheng, Q., et al. (2017). Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proc. Natl. Acad. Sci. U.S.A.* 114, 6848–6853. doi: 10.1073/pnas.1700460114

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Lamouche, Bonadé-Bottino, Mergaert and Alunni. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

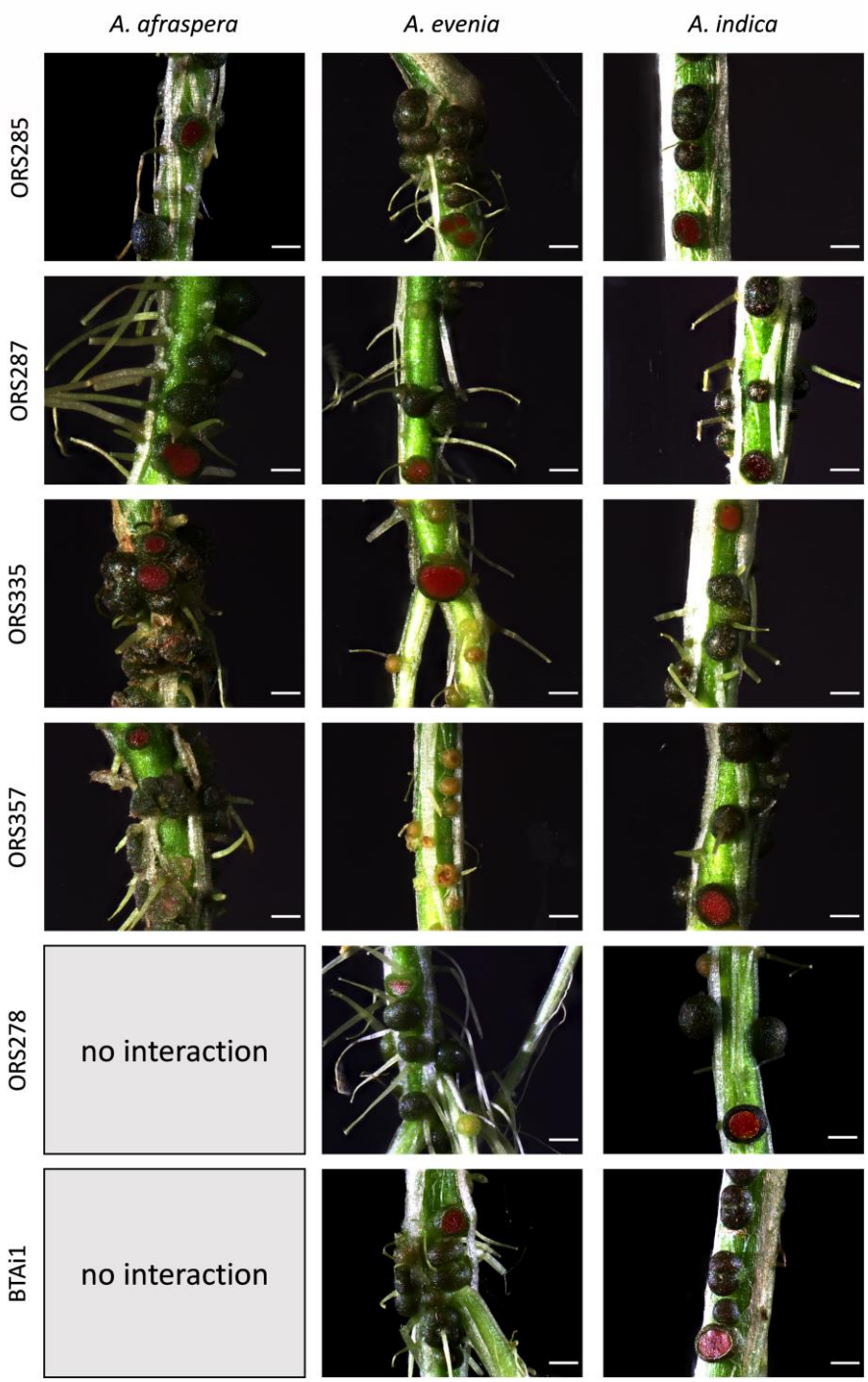


Figure S1. ***Aeschynomene-Bradyrhizobium*** **associations form functional nodules.** Hand-made sections of nodules show the inner part of the symbiotic organ that display a red color indicating that leghemoglobin, a marker of nodule functioning, is produced. Scale bars: 1 mm.

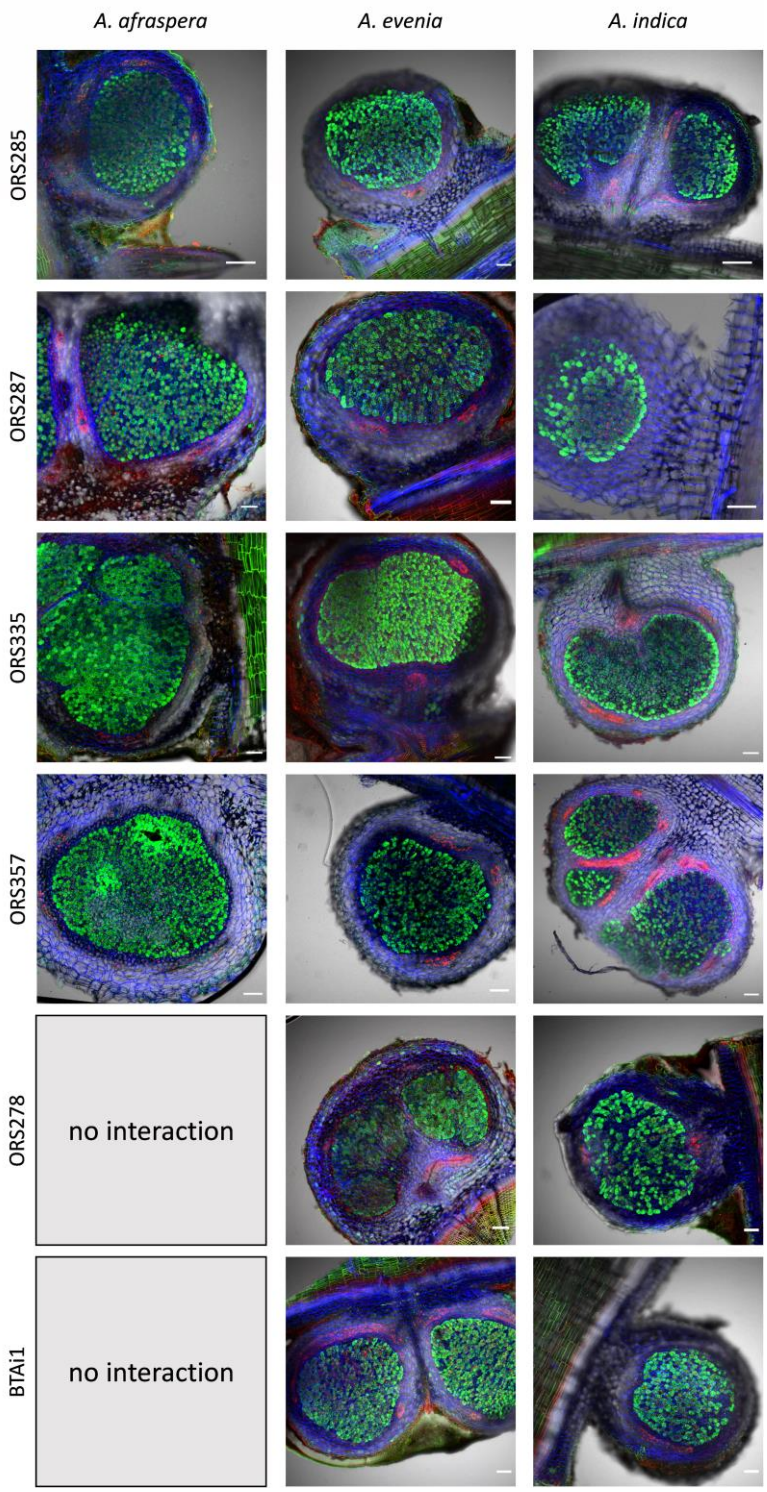


Figure S2. Functional symbiosis is associated with the presence of central nodule tissues composed of symbiotic cells infected by rhizobia. Fourteen dpi nodules were sectioned (70 µm) and stained with LIVE/DEAD kit (live/dead bacteria appear green/red respectively) and calcofluor white (plant cell walls appear blue) before confocal imaging. Scale bars: 100 µm.

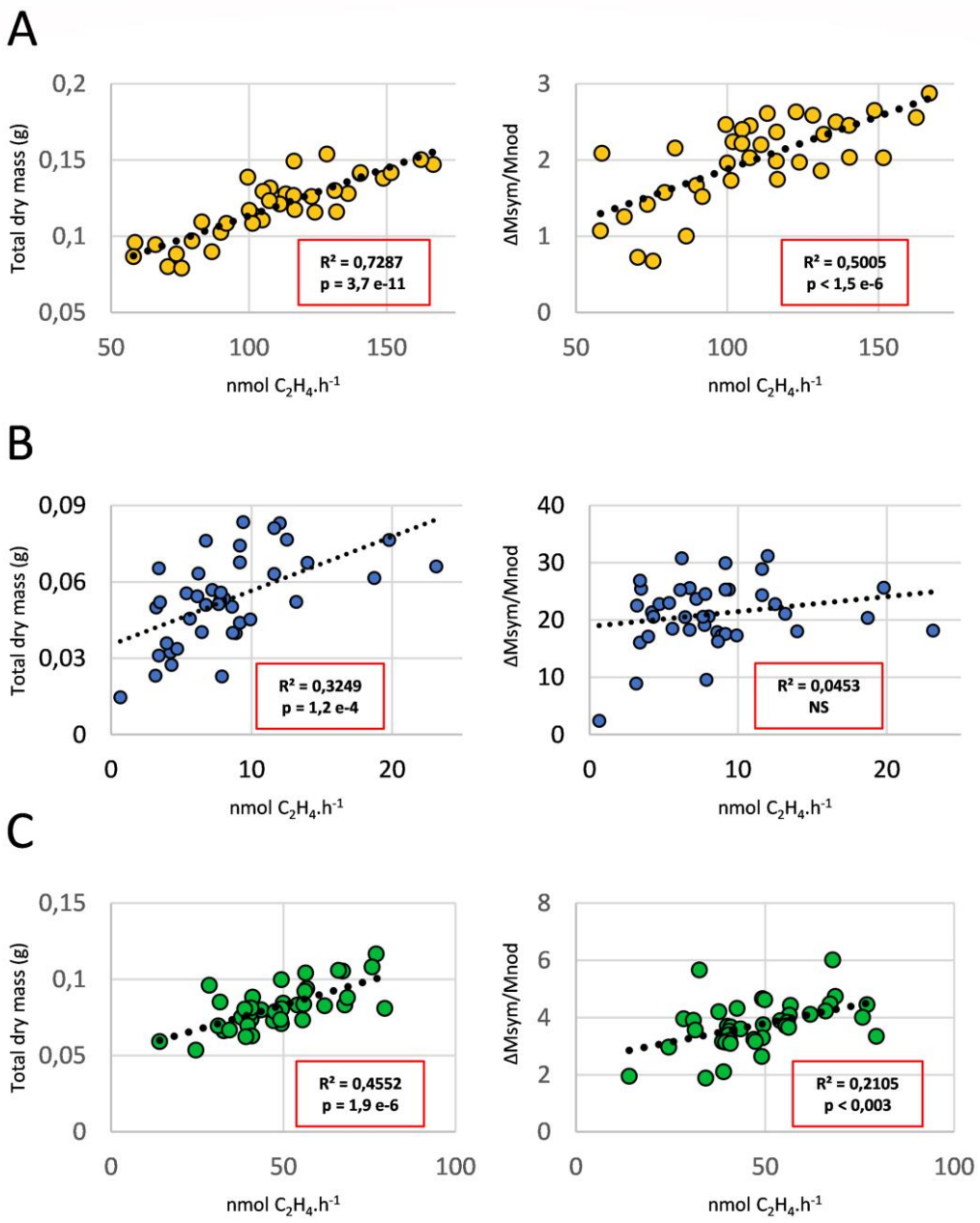


Figure S3. Plant biomass parameters correlate with nitrogen fixation activity in symbiotic systems involving *Bradyrhizobium* sp. ORS285 and *A. afraspera* or *A. indica* and to a lesser extent with *A. evenia*. Relationship between plant total dry mass (left panel) or symbiotic efficiency ($\Delta M_{sym}/M_{nod}$, right panel) and nitrogenase activity (analyzed by acetylene reduction assay and expressed as nmol C_2H_2 evolved per hour and per plant) was tested by calculating the correlation coefficient (R^2) and its associated p-value for 14 dpi nodulated plants of *A. afraspera* (A), *A. evenia* (B) and *A. indica* (C).

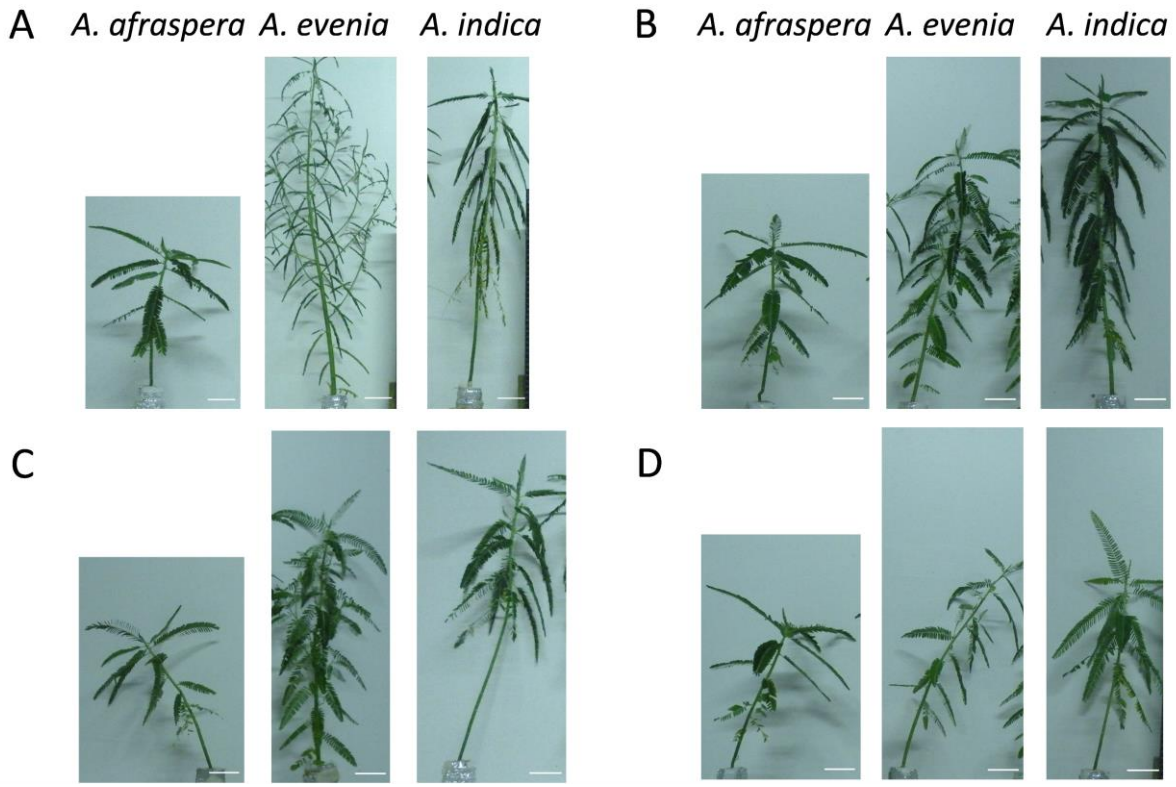


Figure S4. General aspect of *Aeschynomene* plants inducing E- and S-type bacteroids at 45 dpi with ORS285 (A), ORS335 (B), ORS287 (C) or ORS357 (D). The small seed producing *A. evenia* and *A. indica* developed a higher shoot length than the large seed producing *A. afraspera* regardless of the symbiotic strain. Scare bars: 5 cm**.**

2. Transcriptomic dissection of *Bradyrhizobium* sp. strain ORS285 in symbiosis with *Aeschynomene* spp. inducing different bacteroid morphotypes with contrasted symbiotic efficiency

Florian Lamouche, Djamel Gully, Anaïs Chaumeret, Nico Nouwen, Camille Verly, Olivier Pierre, Coline Sciallano, Joël Fardoux, Christian Jeudy, Attila Szücs, Samuel Mondy, Christophe Salon, István Nagy, Attila Kereszt, Yves Dessaux, Eric Giraud, Peter Mergaert et Benoit Alunni

Avant-propos

Les résultats du chapitre 1 ont montré une meilleure efficacité des bactéroïdes ayant un morphotype sphérique. En utilisant la souche *Bradyrhizobium* sp. ORS285, des études transcriptomiques ont été mené pour mieux comprendre la physiologie et le métabolisme des bactéroïdes différenciés dans deux espèces d'*Aeschynomene* induisant des morphotypes allongés ou sphériques. Les potentielles signatures moléculaires responsables de cette différence d'efficacité ont également été investiguées. Trois conditions ont été générées, une culture de référence en milieu riche (YM) et les bactéroïdes de *A. indica* et de *A. afraspera* à 14dpi. Nous avons également décidé d'approfondir l'observation des résultats précédents montrant l'efficacité améliorée des bactéroïdes sphériques par rapport aux bactéroïdes allongés. J'ai ajouté la condition *Aeschynomene nilotica* donnant des bactéroïdes allon gés pour les mesures de masse. Nous nous sommes penchés sur d'autres méthodes, en intégrant des données d'analyse élémentaire (générées par un ancien post-doc, Olivier Pierre) et de fluxomique. J'ai produit les échantillons et j'ai réalisé cette dernière approche en collaboration avec Christian Jeudy et Christophe Salon (INRA Dijon).Le jeu de données de type RNA-seq de la souche sauvage de *Bradyrhizobium* sp. ORS285 analysé dans cette étude a été produit juste avant mon arrivée. J'ai pris en charge le reste du travail, à savoir les étapes de *mapping*, d'analyse différentielle en utilisant le package R DESeq2, de représentation des données sous format publiable et de sélection des gènes d'intérêt avec la littérature associée (Love *et al.*, 2014). J'ai ensuite réalisé les analyses de RT-qPCR afin de confirmer les résultats du transcriptome et de discriminer les gènes dont l'expression est spécifique d'un morphotype ou d'une plante hôte. Des analyses fonctionnelles ont ensuite été faites en parallèle à Gif et dans l'équipe d'Eric Giraud (LSTM, Montpellier). J'ai produit 17 mutants d'insertion chez *Bradyrhizobium* sp. ORS285 et ai testé leur phénotype *in planta*.

Ce travail a été publié dans la revue *Environmental Microbiology* en 2018. DOI : https://doi.org/10.1111/1462-2920.14292

environmental microbiology

Environmental Microbiology (2018) **00**(00), 00–00 doi:10.1111/1462-2920.14292

Transcriptomic dissection of *Bradyrhizobium* sp. strain ORS285 in symbiosis with *Aeschynomene* spp. inducing different bacteroid morphotypes with contrasted symbiotic efficiency

Florian Lamouche,[1†‡] Djamel Gully,[2††] Anaïs Chaumeret,[1†] Nico Nouwen,[2] Camille Verly,[1‡] Olivier Pierre,[1#] Coline Sciallano,[2] Joël Fardoux,[2] Christian Jeudy,[3] Attila Szücs,[4§] Samuel Mondy,[1¶] Christophe Salon,[3] István Nagy,[4,5] Attila Kereszt,[4,5] Yves Dessaux,[1] Eric Giraud,[2] Peter Mergaert[1*] and Benoit Alunni[1*]

[1]*Institute for Integrative Biology of the Cell, UMR 9198, CNRS/Université Paris-Sud/CEA, 91198, Gif-sur-Yvette, France.*

[2]*Laboratoire des Symbioses Tropicales et Méditerranéennes, Institut pour la Recherche et le Développement, UMR IRD/SupAgro/INRA/UM2/CIRAD, Campus International de Baillarguet, TA A-82/J, Montpellier, 34398, France.*

[3]*Agroécologie, AgroSup Dijon, INRA, Université Bourgogne Franche-Comté, Dijon, 21065, France.*

[4]*Biological Research Centre, Hungarian Academy of Sciences, Szeged, 6726, Hungary.*

[5]*Seqomics Biotechnology Ltd, Mórahalom, 6782, Hungary*

Summary

To circumvent the paucity of nitrogen sources in the soil legume plants establish a symbiotic interaction with nitrogen-fixing soil bacteria called rhizobia. During symbiosis, the plants form root organs called nodules, where bacteria are housed intracellularly and become active nitrogen fixers known as bacteroids. Depending on their host plant, bacteroids can adopt different morphotypes, being either unmodified (U), elongated (E) or spherical (S). E- and S-type bacteroids undergo a terminal differentiation leading to irreversible morphological changes and DNA endoreduplication. Previous studies suggest that differentiated bacteroids display an increased symbiotic efficiency (E > U and S > U). In this study, we used a combination of *Aeschynomene* species inducing E- or S-type bacteroids in symbiosis with *Bradyrhizobium* sp. ORS285 to show that S-type bacteroids present a better symbiotic efficiency than E-type bacteroids. We performed a transcriptomic analysis on E- and S-type bacteroids formed by *Aeschynomene afraspera* and *Aeschynomene indica* nodules and identified the bacterial functions activated in bacteroids and specific to each bacteroid type. Extending the expression analysis in E- and S-type bacteroids in other *Aeschynomene* species by qRT-PCR on selected genes from the transcriptome analysis narrowed down the set of bacteroid morphotype-specific genes. Functional analysis of a selected subset of 31 bacteroid-induced or morphotype-specific genes revealed no symbiotic phenotypes in the mutants. This highlights the robustness of the symbiotic program but could also indicate that the bacterial response to the plant environment is partially anticipatory or even maladaptive. Our analysis confirms the correlation between differentiation and efficiency of the bacteroids and provides a framework for the identification of bacterial functions that affect the efficiency of bacteroids.© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd

Received 27 April, 2017; revised 18 May, 2017; accepted 19 May, 2017. *For correspondence. E-mail peter.mergaert@i2bc.paris-saclay.fr; Tel. + 33 1 69 82 34 81. E-mail benoit.alunni@i2bc.paris-saclay.fr; Tel. + 33 1 69 82 34 92. These authors contributed equally to this work.

Present addresses: †Laboratoire de Reproduction et Développement des Plantes, Lyon, France; ‡Institut Jean-Pierre Bourgin, Versailles, France; #Institut Sophia Antipolis, France; §Department of Medical Biology, Szeged University, Szeged, Hungary; ¶Agroécologie, AgroSup Dijon INRA, Université Bourgogne Franche-Comté, Dijon, France. ††These authors contributed equally to this work.

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd.

Introduction

Legume plants circumvent the paucity of soil nitrogen sources by developing symbiotic organs called nodules in which bacteria collectively called rhizobia fix atmospheric nitrogen and transfer ammonia to the plant (Udvardi and Poole, 2013). During nodule formation, rhizobia are released inside the host plant cells by an

endocytosis-like process and transform into nitrogen-fixing bacteroids. In several legume lineages, a terminal bacteroid differentiation (TBD) process occurs after their intracellular uptake, producing elongated polyploid bacterial cells that switched their cell cycle towards endoreduplication (Mergaert *et al*., 2006; Kondorosi *et al*., 2013; Alunni and Gourion, 2016). TBD is controlled by the host plant, which produces defensin-like peptides called Nodule-specific Cysteine Rich (NCR) peptides (Mergaert *et al*., 2003; Van de Velde *et al*., 2010; Czernic *et al*., 2015). NCR peptides are encoded by large gene families. *Medicago truncatula* for example has more than 600 NCR genes (Mergaert *et al*., 2003; Montiel *et al*., 2017). Moreover, the numbers and types of NCR peptides produced by other Inverted-Repeat Lacking Clade (IRLC) legumes correlate with the degree of bacteroid elongation (Montiel *et al*., 2017). In the *Aeschynomene* genus, belonging to the distinct legume clade of the Dalbergioids, the TBD process is also driven by a family of several dozens of NCR peptides, which are however unrelated to the NCR peptides of IRLC and thus arose by a convergent evolution process (Czernic *et al*., 2015).

Legumes of the *Aeschynomene* genus are nodulated by members of the *Bradyrhizobium* genus such as the photosynthetic *Bradyrhizobium* sp. strain ORS285. Depending on the host, bacteroids exhibit different morphotypes, being either elongated (E-type) as in *Aeschynomene afraspera* or *Aeschynomene nilotica*, where bacteroid ploidy raises to 8C (1C being one haploid genome complement), or spherical (S-type) as in *Aeschynomene indica* or *Aeschynomene evenia* where bacteroid ploidy raises to higher levels, up to 16C (Czernic *et al*., 2015; Guefrachi *et al*., 2015). The spherical bacteroid differentiation in *A. indica* comprises a transient cell elongation step (Czernic *et al*., 2015), suggesting that E- and S-type bacteroids can be regarded as a differentiation gradient from less differentiated E-type to more differentiated S-type bacteroids.

The TBD process is not the default state of bacteroids in legumes and several plants do not produce NCR peptides and hence do not induce TBD. For example, soybean induces unmodified bacteroids (U-type) that remain similar to free-living bacteria. Ancestral state reconstruction indicates that TBD is a derived state that evolved independently in different legume clades (Oono *et al*., 2011) in agreement with the independent acquisition of NCR genes in the IRLC and *Aeschynomene*. This convergence toward the evolution of TBD raises questions regarding the selective advantage provided by this process. The biological cost of TBD seems high, as hundreds of peptides are massively produced to induce bacterial differentiation. Studies comparing symbiotic performance of U- *vs.* E-type bacteroids and U- *vs.* S-type bacteroids showed that both types of TBD (E- or S-type nodules) confer a higher symbiotic efficiency than that observed in the absence of TBD (U-type bacteroids) (Sen and Weaver, 1981; Oono and Denison, 2010). These results suggest that TBD increases the efficiency of bacteroids. On the other hand, a comparison of the symbiotic efficiency of E- *vs.* S-type bacteroids has not been performed. Since S-type bacteroids are an enhanced form of differentiation, it is possible that those bacteroids are more efficient than the less differentiated E-type. Moreover, the molecular basis of the increased symbiotic efficiency in E- and S-type bacteroids remains unknown.

In this study, we compared the symbiotic efficiency of E- vs. S-type bacteroids formed by *Aeschynomene* spp. and found an increased symbiotic performance for S-type bacteroids. To identify the underlying basis, a transcriptomic study was conducted. Bacterial functions specifically activated in E- or S-type bacteroids highlighted important metabolic differences between the two bacteroid types. A functional analysis of a selected set of genes identified by transcriptomics, suggested that the bacterial symbiotic program shows robustness and redundancy. This study provides novel insights on the metabolism of differentiated bacteroids and shows to which extent E- and S-morphotype bacteroids behave differently.

Results and discussion

S-morphotype bacteroids have an increased symbiotic efficiency compared to E-morphotype bacteroids

To analyze symbiotic efficiency of E-morphotype and S-morphotype bacteroids, we first determined the plant biomass gained by the symbiosis and per investment in the symbiosis, estimated by the total mass of nodules at 14 dpi (days post inoculation) (see Supporting Information Appendix S1). We find that the S-morphotype inducing plants *A. indica* and *A. evenia*, infected with strain ORS285 obtain a higher benefit than the E-morphotype inducing plants *A. afraspera* and *A. nilotica*, infected with the same strain ORS285 (Fig. 1A). In order to further and more specifically compare the gain in the N-content of the total plant biomass through symbiosis, we analyzed N and C elemental composition of *A. afraspera* and *A. indica* plants infected or not by ORS285 strain at 14 dpi (see Supporting Information Appendix S1). In agreement with the previous result, the S-morphotype inducing plant *A. indica* obtained a higher amount of nitrogen per nodule mass than the E-morphotype inducing plant *A. afraspera* and this was paralleled with a higher increase in carbon as well (Fig. 1B), confirming the different performance of the two bacteroid types. This elemental composition analysis determines the N and C

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

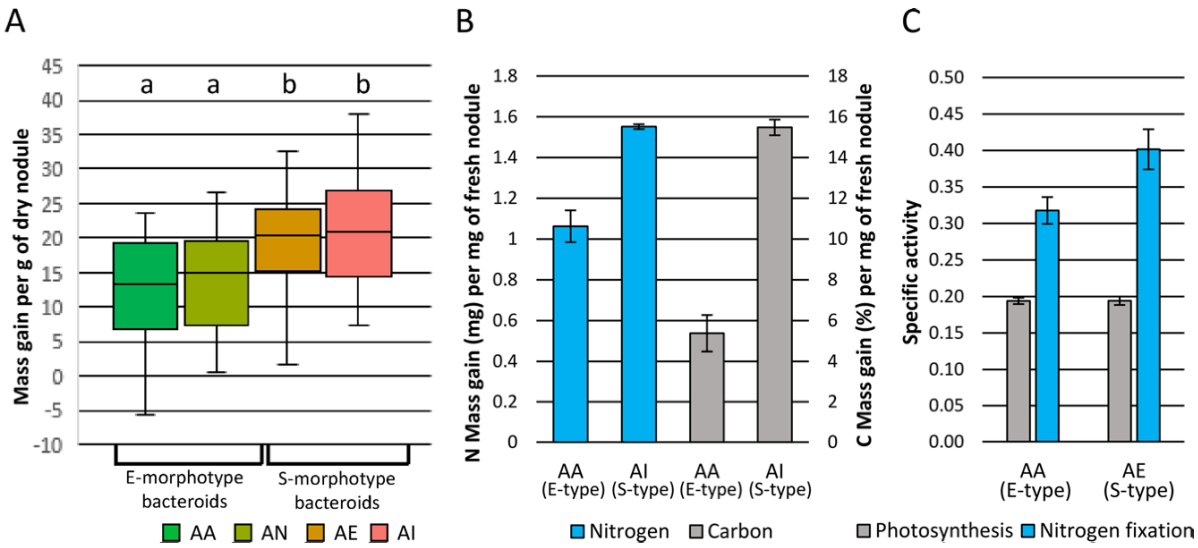


Fig. 1. Evaluation of the symbiotic efficiency of *Bradyrhizobium* sp. ORS285 in symbiosis with *Aeschynomene* spp. plants inducing E- or S-type bacteroids. A. Mass gains at 14 dpi of inoculated plants compared to non-inoculated plants ($n > 30$) and per nodule mass. AA: *Aeschynomene afraspera*; AN: *Aeschynomene nilotica*, AI: *Aeschynomene indica*; AE: *Aeschynomene evenia*. Letters represent significant differences after ANOVA and *post-hoc* Tukey tests ($P < 0.05$). B. Carbon and nitrogen mass gains at 14 dpi determined by elemental analysis of the C and N mass of inoculated versus non-inoculated plants and per nodule mass. C. Specific photosynthetic and N_2 fixation activities from fluxomic analyses at 14 dpi with *Bradyrhizobium* sp. ORS285.

accumulation during the whole lifespan of the plant. Therefore, we analyzed also the N and C fluxes in a 24-hour time window at 14 dpi, a stage of plant growth when nodules are fully developed and have acquired full nitrogen fixation activity (Bonaldi *et al*., 2011). Roots and shoots of *A. evenia* and *A. afraspera* plants at 14dpi were labelled with $^{15}N_2$ and $^{13}CO_2$ isotopes respectively, for one day followed by a 48 h chase and EA-IRMS analysis, to calculate photosynthesis and nitrogen fixation specific activities. Although the photosynthesis specific activity remained comparable for both *A. evenia* and *A. afraspera*, the *Bradyrhizobium* sp. ORS285 provided more nitrogen per nodule mass to *A. evenia* than to *A. afraspera* (Fig. 1C). Together, these results highlight the higher efficiency of S-type bacteroids over E-type bacteroids in *Aeschynomene* spp. plants at 14dpi. To investigate the metabolic processes that could account for this apparent increased symbiotic efficiency we conducted a transcriptomic study by RNA-seq of *Bradyrhizobium* sp. ORS285 grown as free-living cells and in interaction with *A. indica* and *A. afraspera* at 14dpi.

Overview of the transcriptomes

The transcriptomic dataset (Supporting Information Table S1) contained two bacteroid conditions and one culture condition, each in triplicate (Fig. 2A). The sequencing data of this project is available on the NCBI Gene Expression Omnibus (GEO) database with accession n° GSE108744 (SRA number: SRP128034). Principal component analyses revealed a clear partitioning of the samples, the first axis (55% of the observed variance) separating culture from bacteroid conditions and the second axis separating E-morphotype (*A. afraspera*) from S-morphotype (*A. indica*) (20% of the observed variance; Fig. 2B). To assess the validity of the transcriptomic analysis, qRT-PCR was performed on a subset of 23 genes differentially regulated in one of the three conditions and including examples for each of the three conditions (Supporting Information Table S2). The log_2-fold changes (LFC) obtained by both techniques were compared and appeared to be consistent (linear regression R^2 value = 0.77; Fig. 2C).

A first, general analysis of the transcriptomes was obtained by comparing the representation of the different Cluster of Orthologous Genes (COG) under each condition (Fig. 2D). This COG analysis revealed that bacteroids - regardless of their morphotype - undergo a metabolic transition hallmarked by the upregulation of genes involved in energy production/conversion (COG class C) and inorganic ion transport/metabolism (COG class P) on the one hand and by the drastic downregulation of genes involved in translation (COG class J), signal transduction (COG class T), secondary metabolites biosynthesis (COG class Q) and cell motility (COG class N) on the other hand (Fig. 2D). These data corroborate the common view of bacteroid physiology, which is rewired towards nitrogen fixation and export of the fixed nitrogen,

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

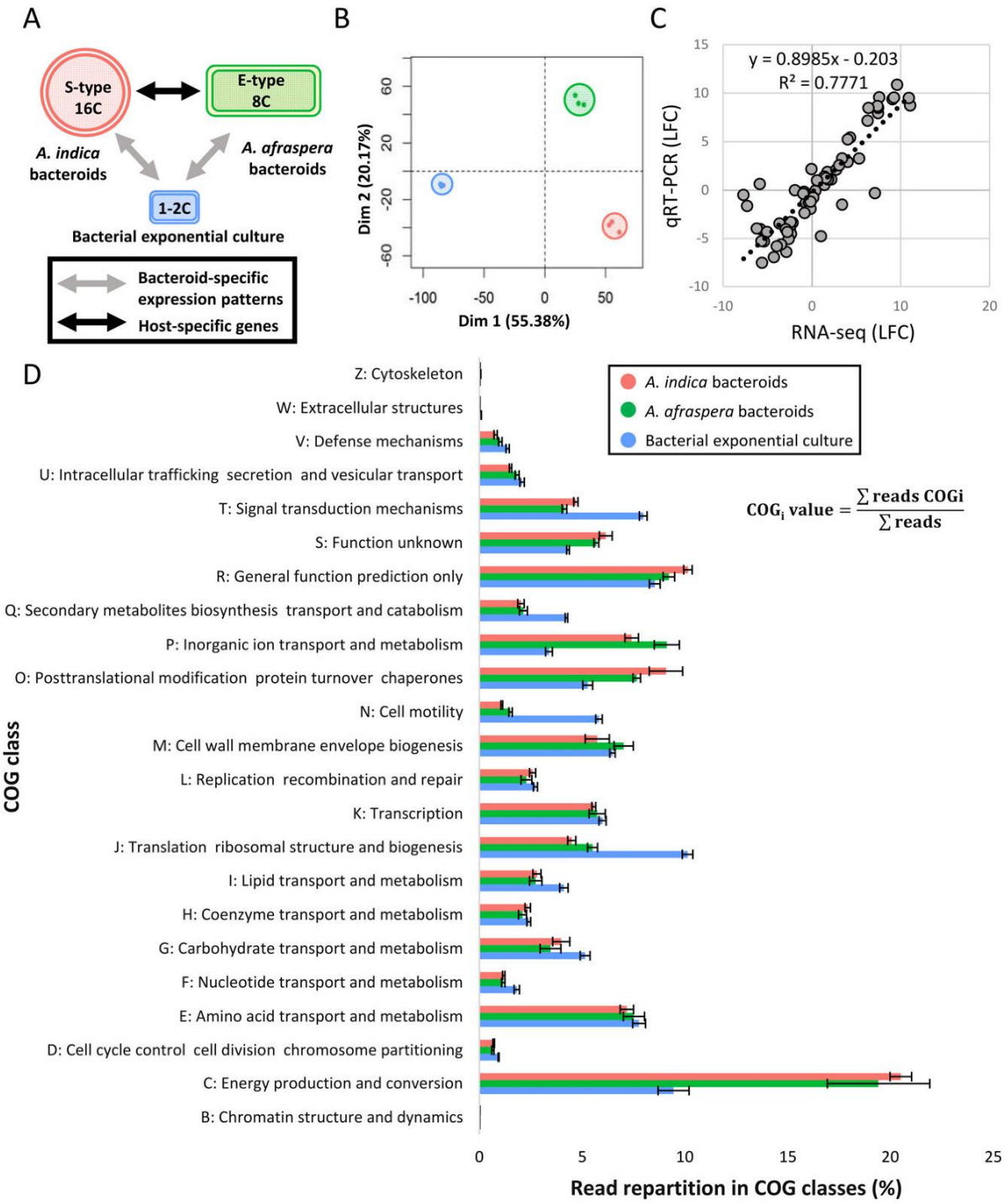


Fig. 2. Experimental setup and general overview of the RNA-seq dataset. A. RNA-seq experimental setup displaying the three biological conditions and the comparisons of interest. B. Principal component analysis. C. qRT-PCR cross validation of RNA-seq results obtained on three independent samples, based on the log2-transformed fold change (LFC) expression values. D. Repartition of the reads among COG classes in the three conditions.

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

whereas several housekeeping functions are shut down, indicating that bacteroids behave as nitrogen-fixing organelles or ammoniaplasts within plant cells (Udvardi and Poole, 2013).

At the gene level, our differential expression analysis, using parameters of False Discovery Rate (FDR) < 0.01 and LFC > 1.58 (fold-change > 3), allowed the identification of 1680 (24.9% of total CDS) and 2123 (31.5%) differentially expressed genes (DEG) for E-bacteroids and S-bacteroids respectively, compared to the YM culture, and 515 (7.6%) DEGs in the E-bacteroid *vs.* S-bacteroid comparison. This number of DEGs between E- and S-type bacteroids is substantially higher as compared to those found by a previous study in *Bradyrhizobium diazoefficiens* on three different phaseoloid host plants, soybean, cowpea and siratro, all forming U-type bacteroids (35, 38 and 129 DEGs per comparison) (Koch *et al.*, 2010). This high number of DEGs between E- and S-type bacteroids is in agreement with a significantly different functioning of these morphotypes.

Common bacteroid-specific genes

The transcriptome dataset contained 1342 DEGs, which are commonly up- or down-regulated in the two bacteroid types (Fig. 3; Supporting Information Table S3). As expected, the genes encoding the nitrogen fixation functions were amongst the most upregulated ones (Supporting Information Fig. S1). Together, the transcripts of this gene cluster represent about 22% of all transcripts in bacteroids. This amazing number illustrates in how far bacteroids are truly ammoniaplasts, converging all their efforts towards nitrogen fixation.

A distinct, strongly upregulated locus carries the genes encoding an uptake hydrogenase and accessory functions (LFC = 2–7; Supporting Information Fig. S2). About one third of the energy consumed by nitrogenase is lost through hydrogen production, which is an obligate by-product of nitrogen fixation. Rhizobia that possess an uptake hydrogenase can recycle this hydrogen and thereby recover some of the lost energy by oxidizing hydrogen. This process is thus a benefit for the plant. Interestingly, the cluster is expressed about twice as high in *A. indica* bacteroids as compared to *A. afraspera* bacteroids. Although this difference does not match the threefold induction (LFC > 1.58) criterion for qualifying these genes as DEGs between the two bacteroid types, this observation suggests a more efficient hydrogen recycling in the spherical bacteroids which could be a contributing factor to the different symbiotic performance of these bacteroids.

The upregulation of phosphate (Supporting Information Fig. S2) and phosphonate (Supporting Information Fig. S1) transport and utilization coupled to the upregulation of a

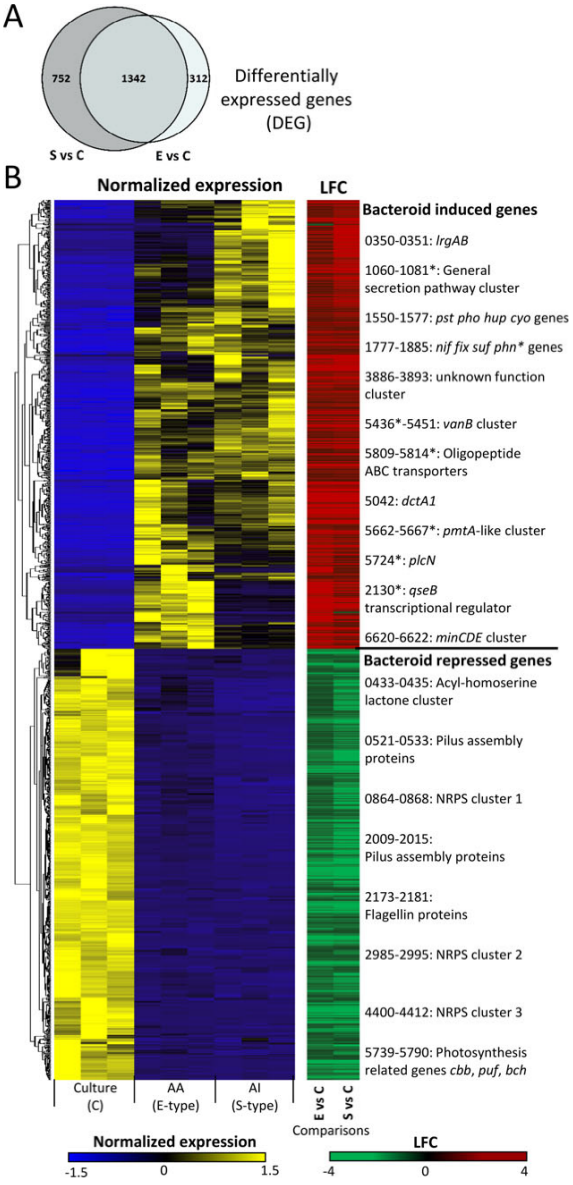


Fig. 3. Overview of the commonly regulated bacteroid gene set. A. Venn diagram showing the overlap of the differentially expressed genes (FDR < 0.01 and |LFC|>1.58) in E- and S-type bacteroids as compared to the culture reference. B. Heatmap and hierarchical clustering of the 1342 differentially expressed genes in both bacteroid comparisons with bacterial culture, displaying bacteroid-specific patterns. The heatmap shows the standard score (z-score) of DESeq2 normalized read numbers. The color-coded scale bars for the normalized expression and LFC of the genes are indicated below the heatmap. Gene accessions are represented by the numerical value of their BRAD285_v2_xxxx format. Asterisks indicate genes for which mutant strains were generated.

set of genes implicated in phospholipid metabolism (Supporting Information Fig. S3) indicate that phosphate and phospholipid homeostasis needs to be modulated in both the E-morphotype and S-morphotype bacteroids

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

(Supporting Information Appendix S2). Remarkably, the locus carrying the phosphonate utilization genes is immediately adjacent to the above-mentioned nitrogen fixation genes (Supporting Information Fig. S1) while the phosphate transport locus is adjacent to the hydrogenase uptake locus (Supporting Information Fig. S2), reinforcing the view that phosphate and phosphonate utilization are important for symbiotic nitrogen fixation. Besides these phosphate and phosphonate transport systems, several other genes involved in metabolite transport were upregulated in the two bacteroid types, reflecting the adaptation of the metabolism of the bacteroids to the intracellular environment in the symbiotic nodule cells (Supporting Information Appendix S2).

Some members of the *Aeschynomene* genus are only effectively nodulated in nature by a Nod factor independent mechanism. They belong to the *Aeschynomene* inoculation group III (Chaintreuil *et al*., 2013). This group is only nodulated by photosynthetic bradyrhizobia and a specific group of non-photosynthetic bradyrhizobia including the strain STM3843 (Okazaki *et al*., 2016; Miché *et al*., 2010). We reasoned that these group of strains may have specific functions that allow them to do so. We used the comparative genomics tool of the MicroScope web-based application (https://www.genoscope.cns.fr; Vallenet *et al*., 2017) to identify gene sets only present in photosynthetic *Aeschynomene* symbionts and absent in other bradyrhizobia. This resulted in the identification of 116 genes (Supporting Information Table S4). Interestingly, we identified several gene regions that were upregulated in bacteroids (Supporting Information Fig. S4). Among them is a large gene cluster of 12 genes (Supporting Information Fig. S5), two other clusters of two or three contiguous genes and several individual genes. Unfortunately, for none of these genes a specific function based on annotation can be proposed (Supporting Information Appendix S2).

Interestingly, a large gene cluster encoding a type II secretion system (T2SS) is strongly upregulated in both bacteroid types (LFC = 2.8–9.6; Supporting Information Fig. S6). T2SSs are widespread in rhizobia but their role in symbiosis has not been analyzed. This machinery secretes proteins to the extracellular environment, ranging from one to several tens of protein substrates depending on the species (Cianciotto and White, 2017). An intriguing possibility is that these T2SS substrates function as (extracellular) effectors in symbiosis in analogy to the effectors secreted by the well-known rhizobial type III secretion systems (T3SS) which suppress host immune responses or induce nodulation (Miwa and Okazaki, 2017). The T2SS locus of strain ORS285 carries in addition to the core components of the apparatus, a few genes that potentially encode secretion substrates. These genes are also strongly or very strongly induced in bacteroids (LFC = 6–9). However, they encode proteins that are only vaguely annotated (see Supporting Information Appendix S2 for a more detailed description of the T2SS in *Bradyrhizobium* sp. ORS285).

Many rhizobia carry their major specific symbiotic functions, including the nitrogenase encoding genes and the nodulation genes for Nod factor biosynthesis, on a plasmid or a symbiotic island (Poole *et al*., 2018). *Bradyrhizobium* sp. strain ORS285 has a symbiotic island carrying the nodulation genes and a T3SS but the above described nitrogenase and hydrogenase genes are not located on it, suggesting that this bacterium may have been a free nitrogen fixer before it became a plant symbiont (see Supporting Information Appendix S2 for a full description of the symbiosis island). We specifically analyzed the expression pattern of the genes located on the island of ORS285 (Supporting Information Fig. S7). None of them showed an upregulation in bacteroids, suggesting that contrary to other rhizobia, the symbiosis island of strain ORS285 has only a function at early stages of nodulation for Nod factor production encoded by the nodulation genes (Nouwen *et al*., 2016). In addition, the T3SS of the symbiosis island in ORS285 does not seem to have a role in bacteroids, in agreement with a previously reported mutant analysis (Okazaki *et al*., 2016).

The two hallmarks of the TBD process are the morphological changes of the bacteria, either elongation to a long rod or inflation to a large sphere, and the polyploidization of their genome (Czernic *et al*., 2015). The nature of these two processes suggests the involvement of the peptidoglycan wall and the cell cycle regulatory machinery in bacteroid differentiation. Accordingly, many of the cell cycle (Supporting Information Fig. S8) and peptidoglycan biosynthesis (Supporting Information Fig. S9) genes show a shift in expression, either up- or down-regulated, in bacteroids confirming the importance of the cell wall dynamics and the cell cycle regulation during bacteroid differentiation (Supporting Information Appendix S2). Other important features involved in the chronic establishment of the bacteroids are the lipopolysaccharide (LPS) and hopanoid lipids of the membrane since mutants that are affected in the production of these lipids form defective bacteroids (Silipo *et al*., 2014; Kulkarni *et al*., 2015; Busset *et al*., 2016, 2017). However, the genes involved in their biosynthesis show no or only slight differences in expression (Supporting Information Figs S10–S12). Only three hopanoid biosynthesis genes *hpnO, hpnP* and *hpnG* are slightly but significantly upregulated in bacteroids suggesting that the bacteroids use hopanoids to adjust the properties of the bacteroid membranes to tune their rigidity in accordance to the local stress condition (see Supporting Information Appendix S2).

Finally, we noticed the strong downregulation of several bacterial functions in bacteroids including pilus formation,

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

motility and chemotaxis (Supporting Information Fig. S13), protein synthesis (Supporting Information Fig. S14), photosynthesis (Supporting Information Fig. S15) and secondary metabolite synthesis (Supporting Information Fig. S16). We refer to Supporting Information Appendix S2 for a further description of these functions.

Identification of morphotype-specific genes

To identify the physiological bases of the differential efficiency of the E- and S-morphotype bacteroids, we next focused on the 515 genes which are differentially expressed between these bacteroid types in the *A. afraspera* and *A. indica* nodules. Hierarchical clustering, including the culture condition, grouped the 515 DEGs into five distinct classes (Fig. 4, Supporting Information Table S5). The first class comprises 142 genes induced only in *A. indica* bacteroids, another class of 136 genes are induced only in *A. afraspera* bacteroids, 35 genes are repressed in *A. afraspera* bacteroids and 35 genes in *A. indica* bacteroids and finally, 167 genes are repressed in both bacteroid types but more strongly in the *A. indica.*

Although a large portion of the DEGs between *A. afraspera* and *A. indica* bacteroids have no or an imprecise annotation (42% are conserved proteins of unknown function and 15% proteins with putative functions), the annotated ones hint at significant difference in the metabolism of both bacteroid types. The most striking difference is the citrate synthase encoding gene *citA* (BRAD285_v2_0083) which is almost 200 times induced in the *A. indica* bacteroids while it is only weakly induced in the *A. afraspera* bacteroids. This high induction makes *citA* having the 11th rank of most highly expressed genes in *A. indica* bacteroids underscoring its importance in these bacteroids. It suggests that in these bacteroids the tricarboxylic acid (TCA) cycle is particularly active. On the other hand, *A. afraspera* bacteroids have a much higher expression of the isocitrate lyase gene *aceA* suggesting that these bacteroids have an enhanced use of the glyoxylate shunt of the TCA cycle as compared to the *A. indica* bacteroids. The higher expression of biotin (*bioADFB*, LFC > 2.1; Supporting Information Fig. S17) and cobalamin (vitamin B_{12}) (16 genes, LFC = 1.3–2.2; Supporting Information Fig. S18) biosynthesis genes in the *A. indica* bacteroids compared to the *A. afraspera* bacteroids is another observation that could hint at a difference in the TCA cycle between these two bacteroid types. These two cofactors are used by the enzymes propionyl-CoA carboxylase and methylmalonyl-CoA mutase (MutB, BRAD285_v2_4653) respectively to form the TCA metabolite succinyl-CoA from propionyl-CoA. We identified in *Bradyrhizobium* sp. strain ORS285 two other enzymes using the cobalamin cofactor. They are methionine synthase (MetH, BRAD285_v2_0497) and the class II ribonucleotide reductase (RNR, BRAD285_v2_3862, *nrd* gene). However while these genes are strongly downregulated in both bacteroid types, the *mutB* gene as well as *mutA* encoding the small subunit of the methylmalonyl-

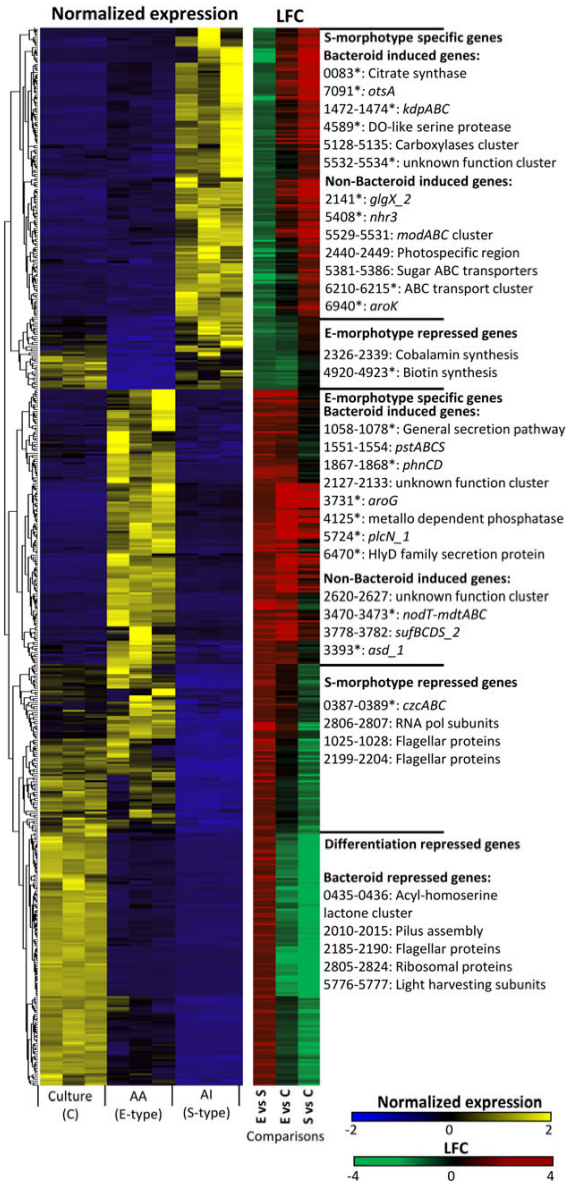


Fig. 4. Heatmap of the 515 differentially expressed genes between E-type and S-type bacteroids. The left heatmap shows the standard score (*z*-score) of the DESeq2 normalized expression of genes in the three experimental conditions and three biological replicates. The heatmap at the right shows the LFC values for pairwise comparisons between the mean of the conditions. The color-coded scale bars for the normalized expression and LFC of the genes are indicated below the heatmap. Gene accessions are represented by the numerical value of their BRAD285_v2_xxxx format. Asterisks indicate genes for which mutant strains were generated.

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

CoA mutase are slightly upregulated, suggesting that this enzyme is active in bacteroids. The two genes encoding the alpha- and beta-subunits of the propionyl-CoA carboxylase are also highly expressed in both bacteroid-types. Thus, the biotin and vitamin B12 production could be increased in the *A. indica* bacteroids in response to a higher demand of the TCA metabolite succinyl-CoA.

In addition, the inferred higher biotin production in the *A. indica* bacteroids is correlated with the higher expression of several other biotin-dependent carboxylases (Supporting Information Fig. S17). This includes three carboxylase subunits encoded by a cluster of seven genes organized in one or two operons (from BRAD285_v2_5128 to BRAD285_v2_5135). The cluster is only vaguely annotated but carries besides the carboxylase genes, genes encoding an ABC transporter as well as enzymes involved in the metabolism of amines suggesting an implication in nitrogen metabolism.

Besides these examples, several other genes differentially expressed between the two bacteroid types hint at metabolic differences. They include for example several amino acid biosynthesis genes. However, pinpointing precise pathways is in many cases difficult because of incomplete annotations. In addition, the many transporter-encoding genes (ABC transporters, outer membrane proteins, RND efflux transporters) among these DEGs also confirm metabolic differences.

A set of DEGs encoding proteins with functions related to bacterial stress responses underlies in another way the differences between the two bacteroid types. Among the most strongly expressed genes in the *A. indica* bacteroids, 18th rank, is *csbD* (BRAD285_v2_5411). It has an almost 20-fold induction while in the *A. afraspera* it is less than threefold induced. A second *csbD* gene, BRAD285_v2_0601, shows a similar pattern but is less strongly expressed. CsbD is widely conserved in bacteria and is a general stress response protein although its precise role remains to be determined (Prágai and Harwood, 2002). Adjacent to the *csbD* gene BRAD285_v2_5411 is the gene BRAD285_v2_5408, also very strongly and specifically induced (about 130 fold) and highly expressed (43th rank) in the *A. indica* bacteroids. This gene encodes a GTPase with unknown function but it is also widely conserved in bacteria and contributes to bacterial fitness in various stress conditions in different bacteria (Khil and Camerini-Otero, 2002; http://fit.genomics.lbl.gov/). We further identified the PhyR-σ^{EcfG} signaling cascade (BRAD285_v2_0553 to BRAD285_v2_0555) as upregulated in bacteroids but two to three times higher in the *A. indica* bacteroids. This cascade is a master regulator of the general stress response mediating resistance to various stresses in Alphaproteobacteria (Gourion *et al*., 2009). Interestingly, a Tn5 mutant in the σ^{EcfG} gene of *Bradyrhizobium* sp. ORS278, a strain closely related to ORS285 formed non-functional nodules on *A. indica* (Bonaldi *et al*., 2010) and corresponding mutants in *Bradyrhizobium japonicum* strain USDA110 similarly formed defective nodules on soybean and mungbean (Gourion *et al*., 2009). The *otsA* gene (BRAD285_v2_7091) encodes the trehalose-phosphate synthase, which catalyzes the first step in the synthesis of the osmo-protectant disaccharide trehalose. The *otsA* gene is strongly induced in both bacteroid types but to a much higher level in the *A. indica* bacteroids. In addition, the *glgX* gene (BRAD285_v2_2141) has a similar expression pattern as *otsA*. GlgX is a glycogen debranching enzyme providing glucose units that is coupled to trehalose synthesis in some bacteria (Schneider *et al*., 2000; Seibold and Eikmanns, 2007). Interestingly, it was reported that overexpression of *otsA* in *Rhizobium etli* resulted in higher nitrogen fixation of *Phaseolus vulgaris* nodules which was correlated with an increased plant biomass, grain-yield and drought tolerance (Suárez *et al*., 2008).

An additional osmo-protectant function specifically activated in the *A. indica* bacteroids is provided by the high-affinity potassium uptake transporter encoded by the *kdpABC* operon (Huang *et al*., 2017). The role of this transporter is to maintain cell shape and turgor. The operon is regulated by the downstream two component regulator genes *kdpDE*. KdpD is a turgor- and osmosensor that is activated by high osmolarity (Jung and Altendorf, 2002). Upon activation of the sensor, the response regulator KdpE is phosphorylated and activates transcription of the *kdpABC* genes. Upstream of *kdpA* are two small open reading frames, encoding peptides of 26 and 29 amino acids, and the gene BRAD285_v2_1469 which are co-transcribed with *kdpABC*. These genes are conserved in *Bradyrhizobium* spp. and could be new components or regulators of this well characterized potassium transporter.

Furthermore, we identified DEGs likely involved in stress regulation that are specifically upregulated in the *A. afraspera* bacteroids. This includes the gene BRAD285_v2_4589, encoding the protease DegP (protease Do). The DegP protease and chaperone is part of the envelope stress response in bacteria and permits the recognition and relief of periplasmic protein-folding stress (Grabowicz and Silhavy, 2017). The genes displaying the highest differential expression in the *A. afraspera* bacteroids compared to the *A. indica* bacteroids, were the *czcABC* genes (LFC > 5) which encode a cation efflux RND transporter that is potentially involved in the regulation of the cation homeostasis. We further identified an additional gene cluster involved in metal homeostasis. Two *suf* clusters are present in the *Bradyrhizobium* sp. ORS285 genome. One of them, *sufBCDSE_1* is part of the nitrogenase locus described above and is induced in both bacteroid types (LFC > 8.8), whereas the second *sufBCDSTA*

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

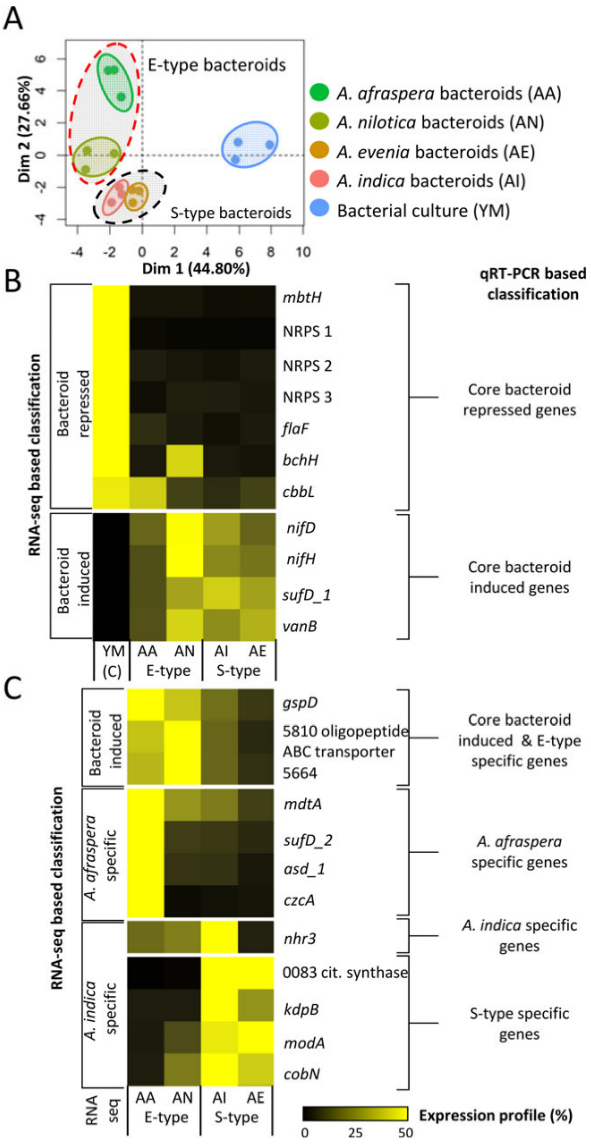


Fig. 5. qRT-PCR expression profiles of selected genes in four plants inducing E-type or S-type bacteroids. A. Principal component analysis of the gene expression profiles in bacterial culture (YM), hosts inducing E-type bacteroids [*A. nilotica* (AN) and *A. afraspera* (AA)] and hosts inducing S-type bacteroids [*A. indica* (AI) and *A. evenia* (AE)] determined by qRT-PCR on biological triplicates for each condition. B and C. Heatmap of the 23 genes of interest measured by qRT-PCR, discriminating bacteroid or culture specificity (B), and morphotype or host specificity (C). Expression levels are shown as percentage of the sum of values for all the samples.

cluster (BRAD285_v2_3776 to BRAD285_v2_3783), located on another genome locus, was specifically expressed in *A. afraspera* bacteroids (LFC > 3.1). The SufABCDSE proteins are involved in iron metabolism by contributing to iron–sulfur cluster biogenesis and repair and they play a critical role in oxidative stress response (Nachin *et al.*, 2003; Jang and Imlay, 2010). It remains to be determined whether the role of the second *sufBCDS* cluster is redundant with that of the first one or specifically necessary for a functional symbiosis in *A. afraspera*.

Taken together our observations suggest that the *A. indica* and *A. afraspera* bacteroids display, besides a number of common responses, also substantial transcriptome differences, notably in metabolism including energy metabolism (hydrogenase, TCA cycle), amino acid and amine metabolism, as well as in the management of stresses imposed by the nodule environment. It will be of importance to determine whether these specificities are related to the different symbiotic efficiencies of the E- and S-morphotype bacteroids in these host plants.

Are differences between A. afraspera *and* A. indica *bacteroids generalizable to E- and S-morphotype bacteroids in other* Aeschynomene *hosts?*

We next addressed the questions whether the observed differential gene expressions between culture and nodule conditions are generalizable to other host plants and whether the observed differential gene expressions between *A. afraspera* and *A. indica* bacteroids are merely related to the host plant or rather reflect fundamental features of the different bacteroid differentiation states. To help answering these two questions, we performed qRT-PCR analyses on cultured bacteria and bacteroids from *A. afraspera* and *A. nilotica* generating E-type bacteroids and *A. evenia* and *A. indica* generating S-type bacteroids. Based on RNA-seq dataset, 23 genes were selected, representing examples of the observed expression profiles. Interestingly, the overall expression profiles of these genes were more similar between hosts inducing the same morphotypes than between hosts inducing contrasted morphotypes, even if the profiles were more variable between the two E-type bacteroids than between the two S-type bacteroids (Fig. 5A). The phylogenetic distance between *A. nilotica* and *A. afraspera* is higher than between *A. evenia* and *A. indica* (Chaintreuil *et al.*, 2013). This could translate into more similar microenvironments for the bacteroids residing within these latter two plants. The tested gene set included 14 genes equally expressed in the *A. afraspera* and *A. indica* transcriptomes, including seven genes repressed in the two bacteroid types and seven upregulated genes. The expression patterns of all these genes were validated by qRT-PCR (Fig. 5B), although three of the bacteroid-enhanced genes seemed to be slightly higher expressed in E-type bacteroids than S-type bacteroids (Fig. 5C). We then tested nine genes specifically induced in *A. indica* or *A. afraspera* respectively, to discriminate if they are host-specific or morphotype-specific. Four genes appeared to be specific

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

of the host *A. afraspera* as they were not upregulated in *A. nilotica* bacteroids. Finally, among the five putative S-morphotype-specific genes tested, one was specific for the *A. indica* host, while the four others appeared to be S-morphotype-specific with similar level of expression in the *A. indica* and *A. evenia* bacteroids and no or weak expression in both E-morphotype bacteroids (Fig. 5C).

Thus overall, the expression patterns revealed by the transcriptome analysis of *A. indica* and *A. afraspera* bacteroids are partially generalizable to S- and E-morphotype bacteroids although it must be recognized that some of the expression profiles are specific for a particular host and are thus not inherent features of particular bacteroid types.

Functional analysis of candidate genes reveals strong robustness and functional redundancy in the symbiotic transcriptome

We selected 31 genes or loci from our transcriptomic dataset that were previously uncharacterized in symbiosis (thus excluding canonical symbiotic genes such as *nif* and *fix* genes) for functional analysis (Supporting Information Table S6). Because of the intriguing upregulation of the T2SS and its complex organization (local duplication of the locus), different deletion mutants were constructed in this locus. Both copies of the T2SS were individually deleted and a double mutant carrying both deletions was also constructed. Furthermore, mutants were constructed in which the single prepilin signal peptidase gene was removed as well as a mutant in which the complete region was deleted. An insertional mutant in one of the T2SS substrate proteins, the putative S-layer protein was also constructed. In addition, insertional mutants in 27 genes were produced (Supporting Information Table S6). The selection criteria were based on their expression profile and level. Indeed, genes chosen are either bacteroid-induced and/or morphotype-specific. Twenty of them have a high expression level in at least one bacteroid condition (first decile of highest expression). Two genes have been selected for their relevance in plant microbe-interactions (BRAD285_v2_5219, putatively involved in salicylic acid metabolism) and nodule development (BRAD285_v2_5527, putatively involved in auxin metabolism).

Mutants were tested for their symbiotic performance. We assessed the capacity of the mutant bacteria to form nodules and sustain plant growth in hosts inducing E-type and S-type bacteroids. Unexpectedly, in all cases, plants displayed pink nodules and green leaves similarly to the controls inoculated with the wild type strain (data not shown). For ten of those mutants and the T2SS deletion mutants, we determined the nodule number per plant and measured nitrogenase activity by acetylene reduction assays. No difference in nodule formation and its kinetics, nodule tissue organization and nitrogen fixation could be identified between plants inoculated by wild-type and mutant bacteria (Supporting Information Figs S19 and S20). This result indicates that none of the candidate genes is essential for symbiosis under the tested conditions, which were identical to the conditions used for the transcriptome analysis. This observation is consistent with a previous report in another symbiotic system, the interaction of *Rhizobium leguminosarum* with pea, where the mutation of 37 bacteroid-induced genes resulted in no detectable symbiotic phenotype (Karunakaran *et al*., 2009).

To further analyze these surprising results, we compiled the expression patterns of all reported genes with symbiotic phenotypes in the corresponding mutants of *Bradyrhizobium* sp. ORS285 or related photosynthetic bradyrhizobia (Supporting Information Table S7). For a total of 131 mutants, the corresponding genes in 93 (71%) cases were not differentially expressed in bacteroids. Only 22 genes (17%) with symbiotic phenotypes were upregulated in bacteroids, seven of them encoding components or regulators of the nitrogenase. Counterintuitively, 16 (12%) of the genes displaying a nodulation phenotype when mutated were even downregulated in the bacteroids. Since the large majority of these mutants were obtained by unbiased forward genetic screens (Giraud *et al*., 2007; Bonaldi *et al*., 2010), we can conclude that genes that result in a symbiotic phenotype when mutated, are not necessarily differentially expressed.

The absence of symbiotic phenotypes in bacteroid-upregulated genes in this and previous work requires further explanation because induced and high expression is in general taken as an adaptive response of the organism or cell to a specific condition so that genes are expressed when they are required. First, it may suggest a strong robustness and a functional redundancy within the bacterial symbiotic program. Robustness implies that each of the induced gene functions have a slight additive contribution to the symbiotic process but that the removal of an individual function has no or an imperceptible effect, not enough to let the edifice crumble, especially under controlled laboratory conditions. For example, the putative T2SS determined S-layer could together with the LPS and the hopanoid lipids constitute a multilayered envelope which provides sufficient strength to the bacteria to withstand the NCR peptides or other stresses in the symbiotic nodule cells. Removing the individual components of this envelope could have an imperceptible or moderate effect on the bacteroids (Silipo *et al*., 2014; Kulkarni *et al*., 2015; Busset *et al*., 2016, 2017; this work) but, possibly, removing multiple parts simultaneously would be much more deleterious for the bacteroids. Redundancy

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd.,

on the other hand can be achieved by homologous genes or through a functional complementation by an unrelated pathway. For some of the insertional mutants, we cannot exclude that redundancy with a homologous gene is responsible for the lack of symbiotic phenotype although none of the selected genes has a homolog of more than 55% (Supporting Information Table S6). The robustness and redundancy could be experimentally verified by the construction of mutants carrying multiple mutations.

It is also possible that the mutated genes have a role in symbiosis under particular conditions that are not captured by our laboratory nodulation and nitrogen fixation assays but that might exist in natural growth conditions. In that case, one has to assume that these genes are not finely regulated but activated to a standby mode even if not needed in the present condition but in such a way that their gene products are available in the advent of a changing condition. In the same vein, the genes could be activated to function in a future stage of the nodule life-cycle such as for example in nodule senescence. Nodule senescence is the last stage of the nodule lifetime and leads to the digestion of the host tissues as well as a large part of the bacteroids. This process is mediated by the massive activation of macromolecule-degrading enzymes by the host (Van de Velde *et al*., 2006). It is possible that bacteroids activate functions that could help them in their attempt to survive the future attack of the host to digest them. Still another possibility is that these genes are only required for symbiosis with certain *Aeschynomene* hosts but not with others while their regulation does not distinguish between differences in host environments. A striking example is the *nifV* gene encoding a homocitrate synthase. We find that this gene is induced to very high and equal levels in both nodule types (LFC > 11; Supporting Information Table S1) but mutagenesis has shown that the gene is only required for symbiosis with *A. indica* and not with *A. afraspera* (Nouwen *et al*., 2017). Homocitrate is an essential intermediate in the production of the iron-molybdenum cofactor (FeMo-co) of nitrogenase. Thus, nitrogen fixation in bacteroids requires a functional *nifV* gene except if homocitrate is supplied by the host plant through the nodule-specific expression of the *FEN1* gene, a plant homocitrate synthase gene (Hakoyama *et al*., 2009; Nouwen *et al*., 2017). *A. afraspera* has such a nodule-specific *FEN1* gene but *A. indica* not (Nouwen *et al*., 2017). Thus, *Bradyrhizobium* sp. ORS285 activates *nifV* expression in nodules irrespectively of the supply of homocitrate by the host plant.

Finally, we cannot exclude that part of the transcriptome changes that we have observed are non-adaptive responses. From recent findings, the picture emerges that transcriptional responses in bacteria or yeast are in large part suboptimal. Systematic comparisons of mutant fitness and gene expression have shown that these two parameters are poorly correlated (Giaever *et al*., 2002; Deutschbauer *et al*., 2011; Price *et al*., 2013). Thus, genes may be upregulated in a specific context only as a side effect without providing a functional contribution. In the case of rhizobia, the bacteria have a dual lifestyle as free-living soil and rhizosphere bacteria and as nodule symbionts. It is possible that these bacteria have certain adaptations enhancing their fitness during free-living growth in the proximity of plants and that they therefore recognize certain plant-derived signals that induce the genes encoding these adaptations. It is further possible that the same or similar signaling molecules are also present inside nodules thereby enhancing the expression of the same gene set which make however no specific contribution to the symbiosis.

Conclusions

Bacteroids in legume nodules come in different flavors, characterized by their typical size, morphology, degree of polyploidy and membrane permeabilization. The functional implications on these differences are still mostly unclear. Since the differentiation process is imposed by the plant through the production or not of the NCR peptides, it was predicted that the differentiation process provides advantage to the host and thus different efficiencies of these bacteroids were suggested (Oono and Denison, 2010; Kereszt *et al*., 2011). Our analysis, a case study comparing E- and S-morphotype bacteroids formed by two closely related host plants, revealed that S-morphotype bacteroids tended to be more efficient nitrogen fixers than elongated ones at 14dpi. At the molecular level, our transcriptomic survey of differentiated bacteroids displaying contrasted morphotypes allowed the identification of a bacteroid gene set that included determinants of the main metabolic pathways required for nitrogen fixation and symbiotic life. When comparing E- and S-morphotype bacteroids, we identified large gene sets that were specific for one or the other bacteroid type and that could provide hints on candidate functions associated with the increased symbiotic efficiency of spherical bacteroids. Nevertheless, as the two host plants are different species, we have to acknowledge that the detected differences in symbiotic efficiency and in gene expression can only be correlated with bacteroid types and that at this stage, no causal relationship can be concluded. Indeed, as our qRT-PCR data demonstrates, some of the expression differences are not related to the bacteroid type but probably to some other, yet unknown, differences between the hosts. Extending the analyses to a larger number of hosts inducing E- and

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

S-morphotype bacteroids would increase our capacity to distinguish between these confounding factors.

Disappointingly, mutants in a selected set of upregulated genes did not reveal any phenotype even if they were tested under identical conditions as the expression analysis. To go further in the characterization of these genes, possible strategies are enlarging the set of conditions in which phenotypes of mutants are tested or the creation of double, triple… mutants.

Finally, the present analysis could be usefully complemented with a metabolome and proteome comparison of the bacteroids. Our analysis hints at a different metabolism in the E- and S-morphotype bacteroids based on differential gene expression. However, cellular metabolic fluxes are determined more by substrate concentrations than protein abundances (Hackett *et al*., 2016) and protein abundance does not always correlate with gene expression. Thus, metabolome and proteome studies have the potential to further deepen our understanding of the differences between the bacteroid types.

Experimental procedures

Plant growth and inoculation

Bradyrhizobium sp. strain ORS285 was grown at 30 °C in YM medium supplemented with carbenicillin (50 μg ml^{-1}) (Guefrachi *et al*., 2015). *A. indica, A. nilotica, A. evenia* and *A. afraspera* plants were grown in transparent test tubes filled with BNM medium and inoculated with strain ORS285 as described before (Guefrachi *et al*., 2015). Plants were grown at 28 °C, 80% humidity and under a 16h: 8 h of light: dark regimen.

Symbiotic efficiency analyses

The symbiotic efficiencies of the *A. indica, A. nilotica, A. evenia* and *A. afraspera* interactions with *Bradyrhizobium* sp. strain ORS285 were determined at 14 dpi using three independent methods: (i) plant biomass determination, (ii) elemental analysis to determine the nitrogen content in whole plants and (iii) fluxomic analysis to determine the transfer of fixed nitrogen to the plant in 24 h. Details of these methods are provided in Supporting Information Appendix S1.

Transcriptome and qRT-PCR analyses

For gene expression analyses, nodules were harvested at 14 dpi. Procedures for qRT-PCR measurements, RNA extraction, RNA-seq with the SOLiD technology, sequencing data treatment, statistical analyses and genome mining are described in detail in Supporting Information Appendix S3. The DESeq2 (Love *et al*., 2014) normalized RNA-seq data are represented in the heat maps as the standard score (z-score) of the normalized read numbers (Figs 3B and 4) or as the percentage of the sum of the normalized read numbers for all the samples (Supporting Information Figs S1–S18). The qRT-PCR data are shown as the percentage of the sum of values for all the samples (Fig. 5B and C).

Generation and phenotyping of bacterial mutants

For insertion mutagenesis, internal fragments (300-600 bp) of the target gene were PCR-amplified and cloned into the pVO155-nptII-aadA-GFP vector. The resulting constructs were introduced into *Bradyrhizobium* sp. ORS285 and mutants were selected on 50 μg l^{-1} kanamycin and 50 μg ml^{-1} spectinomycin plates (Wongdee *et al*., 2016). For the construction of deletion mutants, flanking regions of the gene were PCR amplified, fused by overlap extension PCR and cloned in the *sacB* suicide pNPTS129 plasmid. The resulting plasmids were introduced into *Bradyrhizobium* sp. ORS285 and mutants were obtained after a double recombination event as described (Nouwen *et al*., 2016). The genotype of all mutant strains was verified by PCR analysis. Primers for constructions and strain verification are listed in Supporting Information Table S8. The phenotype of the mutants was analyzed by inoculating *A. indica* 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).

Acknowledgments

F.L. was 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, grant no ANR-13-BSV7-0013 and used resources from the National Office for Research, Development and Innovation of Hungary, grant no 120120 to A.K.

References

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.

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

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**: 9063.

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.

Bonaldi, K., Gourion, B., Fardoux, J., Hannibal, L., Cartieaux, F., Boursot, M., *et al*. (2010) Large-scale transposon mutagenesis of photosynthetic *Bradyrhizobium* sp. strain ORS278 reveals new genetic loci putatively important for nod-independent symbiosis with *Aeschynomene indica*. *Mol Plant Microbe Interact* **23**: 760–770.

Busset, N., De Felice, A., Chaintreuil, C., Gully, D., Fardoux, J., Romdhane, S., *et al*. (2016) The LPS O-antigen in photosynthetic *Bradyrhizobium* strains is dispensable for the establishment of a successful symbiosis with *Aeschynomene* legumes. *PLoS One* **11**: e0148884.

Busset, N., Di Lorenzo, F., Palmigiano, A., Sturiale, L., Gressent, F., Fardoux, J., *et al*. (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. *Front Microbiol* **8**: 1821.

Chaintreuil, C., Arrighi, J.-F., Giraud, E., Miché, L., Moulin, L., Dreyfus, B., *et al*. (2013) Evolution of symbiosis in the legume genus *Aeschynomene*. *New Phytol* **200**: 1247–1259.

Cianciotto, N.P., and White, R.C. (2017) Expanding role of type II secretion in bacterial pathogenesis and beyond. *Infect Immun* **85**: 11–16.

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., *et al*. (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.

Deutschbauer, A., Price, M.N., Wetmore, K.M., Shao, W., Baumohl, J.K., Xu, Z., *et al*. (2011) Evidence-based annotation of gene function in *Shewanella oneidensis* MR-1 using genome-wide fitness profiling across 121 conditions. *PLoS Genet* **7**: e1002385.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., *et al*. (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**: 387–391.

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.-C., *et al*. (2007) Legumes symbioses: absence of nod genes in photosynthetic bradyrhizobia. *Science* **316**: 1307–1312.

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.

Grabowicz, M., and Silhavy, T.J. (2017) Redefining the essential trafficking pathway for outer membrane lipoproteins. *Proc Natl Acad Sci U S A* **114**: 4769–4774.

Guefrachi, I., Pierre, O., Timchenko, T., Alunni, B., Barriere, 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.

Hackett, S.R., Zanotelli, V.R.T., Xu, W., Goya, J., Park, J.O., Perlman, D.H., *et al*. (2016) Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science* **354**: aaf2786.

Hakoyama, T., Niimi, K., Watanabe, H., Tabata, R., Matsubara, J., Sato, S., *et al*. (2009) Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. *Nature* **462**: 514–517.

Huang, C.-S., Pedersen, B.P., and Stokes, D.L. (2017) Crystal structure of the potassium-importing KdpFABC membrane complex. *Nature* **546**: 681–685.

Jang, S., and Imlay, J.A. (2010) Hydrogen peroxide inactivates the *Escherichia coli* Isc iron-sulphur assembly system, and OxyR induces the Suf system to compensate. *Mol Microbiol* **78**: 1448–1467.

Jung, K., and Altendorf, K. (2002) Towards an understanding of the molecular mechanisms of stimulus perception and signal transduction by the KdpD/KdpE system of *Escherichia coli*. *J Mol Microbiol Biotechnol* **4**: 223–228.

Karunakaran, R., Ramachandran, V.K., Seaman, J.C., East, A.K., Mouhsine, B., Mauchline, T.H., *et al*. (2009) Transcriptomic analysis of *Rhizobium leguminosarum* biovar *viciae* in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. *J Bacteriol* **191**: 4002–4014.

Kereszt, A., Mergaert, P., and Kondorosi, E. (2011) Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims?. *Mol Plant Microbe Interact* **24**: 1300–1309.

Khil, P.P., and Camerini-Otero, R.D. (2002) Over 1000 genes are involved in the DNA damage response of *Escherichia coli*. *Mol Microbiol* **44**: 89–105.

Koch, M., Delmotte, N., Rehrauer, H., Vorholt, J., 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.

Kondorosi, E., Mergaert, P., and Kereszt, A. (2013) A paradigm for endosymbiotic life: cell differentiation of Rhizobium bacteria provoked by host plant factors. *Annu Rev Microbiol* **67**: 611–628.

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**: e01251-15.

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

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**: 161–173.

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., *et al*. (2006) Eukaryotic control on bacterial cell

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

cycle and differentiation in the Rhizobium-legume symbiosis. *Proc Natl Acad Sci U S A* **103**: 5230–5235.
Miché, L., Moulin, L., Chaintreuil, C., Contreras-Jimenez, J. L., Munive-Hernández, J., Del Carmen Villegas-Hernandez, M., *et al*. (2010) Diversity analyses of *Aeschynomene* symbionts in Tropical Africa and Central America reveal that *nod*-independent stem nodulation is not restricted to photosynthetic bradyrhizobia. *Environ Microbiol* **12**: 2152–2164.
Miwa, H., and Okazaki, S. (2017) How effectors promote beneficial interactions. *Curr Opin Plant Biol* **38**: 148–154.
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.
Nachin, L., Loiseau, L., Expert, D., and Barras, F. (2003) SufC: an unorthodox cytoplasmic ABC/ATPase required for [Fe-S] biogenesis under oxidative stress. *Embo J* **22**: 427–437.
Nouwen, N., Arrighi, J.-F., Cartieaux, F., Chaintreuil, C., Gully, D., Klopp, C., *et al*. (2017) The role of rhizobial (NifV) and plant (FEN1) homocitrate synthases in *Aeschynomene*/photosynthetic *Bradyrhizobium* symbiosis. *Sci Rep* **7**: 448.
Nouwen, N., Fardoux, J., and Giraud, E. (2016) NodD1 and NodD2 are not required for the symbiotic interaction of *Bradyrhizobium* ORS285 with nod-factor-independent *Aeschynomene* legumes. *PLoS ONE* **11**: e0157888.
Okazaki, S., Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., *et al*. (2016) Rhizobium-legume symbiosis in the absence of Nod factors: Two possible scenarios with or without the T3SS. *ISME J* **10**: 64–74.
Oono, R., Anderson, C.G., and Denison, R.F. (2011) Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. *Proc R Soc Lond B Biol Sci* **278**: 2698–2703.
Oono, R., and Denison, R.F. (2010) Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol* **154**: 1541–1548.
Poole, P., Ramachandran, V., and Terpolilli, J. (2018) Rhizobia: from saprophytes to endosymbionts. *Nat Rev Microbiol* **16**: 291–303.
Prágai, Z., and Harwood, C.R. (2002) Regulatory interactions between the Pho and sigma(B)-dependent general stress regulons of *Bacillus subtilis*. *Microbiology* **148**: 1593–1602.
Price, M.N., Deutschbauer, A.M., Skerker, J.M., Wetmore, K. M., Ruths, T., Mar, J.S., *et al*. (2013) Indirect and suboptimal control of gene expression is widespread in bacteria. *Mol Syst Biol* **9**: 660.
Schneider, D., Bruton, C.J., and Chater, K.F. (2000) Duplicated gene clusters suggest an interplay of glycogen and trehalose metabolism during sequential stages of aerial mycelium development in *Streptomyces coelicolor* A3(2). *Mol Gen Genet* **263**: 543–553.
Seibold, G.M., and Eikmanns, B.J. (2007) The *glgX* gene product of *Corynebacterium glutamicum* is required for glycogen degradation and for fast adaptation to hyperosmotic stress. *Microbiology* **153**: 2212–2220.
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.
Silipo, A., Vitiello, G., Gully, D., Sturiale, L., Chaintreuil, C., Fardoux, J., *et al*. (2014) Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. *Nat Commun* **5**: 5106.
Suárez, R., Wong, A., Ramírez, M., Barraza, A., Orozco, M. D C., Cevallos, M.A., *et al*. (2008) Improvement of drought tolerance and grain yield in common bean by overexpressing trehalose-6-phosphate synthase in rhizobia. *Mol Plant Microbe Interact* **21**: 958–966.
Udvardi, M., and Poole, P.S. (2013) Transport and metabolism in legume-rhizobia symbioses. *Annu Rev Plant Biol* **64**: 781–805.
Vallenet, D., Calteau, A., Cruveiller, S., Gachet, M., Lajus, A., Josso, A., *et al*. (2017) MicroScope in 2017: An expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res* **45**: D517–D528.
Van de Velde, W., Guerra, J.C.P., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., *et al*. (2006) Aging in legume symbiosis. A molecular view on nodule senescence in *Medicago truncatula*. *Plant Physiol* **141**: 711–720.
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.
Wongdee, J., Songwattana, P., Nouwen, N., Noisangiam, R., Fardoux, J., Chaintreuil, C.E., *et al*. (2016) *nifDK* clusters located on the chromosome and megaplasmid of *Bradyrhizobium* sp. strain DOA9 contribute differently to nitrogenase activity during symbiosis and free-living growth. *Mol Plant Microbe Interact* **29**: 767–773.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Methods for symbiotic efficiency analyses.

Appendix S2. Complementary results on common bacteroid-specific genes

Appendix S3. Methods for transcriptome analysis.

Table S1. Full dataset of *Bradyrhizobium* sp. ORS285 transcriptome.

Table S2. Genes used for qRT-PCR experiments.

Table S3. Dataset of bacteroid-specific genes.

Table S4. Common genes specific for photosynthetic, Nod-independent *Aeschynomene* symbionts.

Table S5. Bacteroid morphotype-specific genes.

Table S6. List of constructed *Bradyrhizobium* sp. ORS285 mutants.

Table S7. Expression of genes with published mutant symbiotic phenotypes in strain ORS285 or related photosynthetic bradyrhizobia.

Table S8. Primers used for qRT-PCR on *Bradyrhizobium* sp. ORS285 and for mutagenesis.

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

Fig. S1. Nitrogenase region of *Bradyrhizobium* sp. ORS285 is highly induced in bacteroids.

Fig. S2. The *Bradyrhizobium* sp. ORS285 locus containing *pst* and *hup* genes is induced in bacteroid conditions.

Fig. S3. *Bradyrhizobium* sp. ORS285 bacteroids highly express a genomic region involved in phospholipid metabolism as well as phospholipases C.

Fig. S4. Expression of photosynthetic *Bradyrhizobium*-specific genes.

Fig. S5. Expression in *Bradyrhizobium* sp. ORS285 of the photosynthetic *Bradyrhizobium*-specific *vanB* gene cluster.

Fig. S6. *Bradyrhizobium* sp. ORS285 features a duplicated Type II Secretion System (T2SS) induced in bacteroids.

Fig. S7. Expression of the *Bradyrhizobium* sp. ORS285 symbiotic island.

Fig. S8. Expression of cell cycle regulatory genes.

Fig. S9. Expression of peptidoglycan biosynthesis genes.

Fig. S10. Expression of the hopanoid biosynthesis pathway.

Fig. S11. Expression of the lipopolysaccharide (LPS) biosynthesis pathway.

Fig. S12. Expression of polysaccharide biosynthesis genes.

Fig. S13. Expression of genes involved in chemotaxis, flagellum and pilus biosynthesis.

Fig. S14. Expression of ribosomal genes.

Fig. S15. Expression of genes involved in bacterial photosynthesis.

Fig. S16. Expression of Non-Ribosomal Peptide Synthase (NRPS) and Quorum Sensing (QS) genes.

Fig. S17. Expression of genes involved in biotin biosynthesis and utilization.

Fig. S18. Expression of genes involved in vitamin B12 synthesis and utilization.

Fig. S19. Nodulation phenotypes of eight mutant strains.

Fig. S20. Nodulation phenotypes of mutants in the general secretion pathway.

© 2018 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

Supporting information

Appendix S1. Methods for symbiotic efficiency analyses.

Plant biomass production

Plants (n≥30) were collected at 14 days post inoculation (dpi). Nodules, roots and shoots were separated and dried at 80°C for 48h after nodules have been counted. Dry mass was measured for the shoot, root and nodule plant compartments. The proxy used to measure symbiotic efficiency was the total plant dry mass gain per mg of nodule, calculated as the dry mass difference between the inoculated and non-inoculated plants, divided by the nodule dry mass. Significant differences were determined by analysis of variance followed by a *post hoc* HSD Tukey test ($p<0.05$).

Elemental analysis

The nitrogen (N) and carbon (C) content was determined as % (m/m) on pools of three de-nodulated 14 dpi whole plants that were dried at 60 °C for 48 h. The dried samples were analyzed with an Isoprime element analyzer (Elementar). Mass gain per mg of nodule has been estimated as a proxy of symbiotic efficiency.

Fluxomic analysis

Forty *A. afraspera* and *A. evenia* plants were grown until 14 dpi under the conditions described above. Roots and shoots were placed into hermetically sealed devices. Root and shoot compartments were then exposed to an enrichment in $^{15}N_2$ (5%) and $^{13}CO_2$ (30%), respectively, for 24h followed by a 48h flush. Plants were harvested and separated into aerial parts, roots and nodules. Samples were then dried and crushed. For each sample, ca. 2 mg of dry powder was analyzed by elemental analyzer coupled to isotopic ratio mass spectroscopy (EA-IRMS) to quantify the isotopic composition of C and N. Ten *A. evenia* and 2 or 3 *A. afraspera* plants have been pooled to reach the detection threshold of the spectrometer. New biomass measured from $^{15}N_2/^{13}CO_2$ labeling was calculated as described (Voisin *et al.*, 2003). Photosynthesis specific activity and nitrogen-fixation specific activity have been calculated as follows: Photosynthesis

specific activity = Total new C biomass / Aerial parts biomass; Nitrogen-fixation specific activity = Total new N biomass / Nodule biomass.

Reference

Voisin, A.S., Salon, C., Jeudy, C., Warembourg, F.R. (2003) Symbiotic N_2 fixation activity in relation to C economy of *Pisum sativum* L. as a function of plant phenology. *J Exp Bot* **54**: 2733–2744.

Appendix S2. Common bacteroid-specific genes

Phosphate and phosphonate utilization

A set of neighboring genes probably implicated in phospholipid metabolism was upregulated under bacteroid conditions. Three genes involved in phospholipid synthesis were induced (LFC=3.6-8.4; Fig S5), including *psd* and *pmtA*-like *genes*, encoding a phosphatidylserine decarboxylase and a phosphatidylcholine methyltransferase enzyme, respectively. Interestingly, the *pmtA* gene involved in phosphatidylcholine synthesis has been previously reported to be essential for effective nitrogen fixation in the *Bradyrhizobium diazoefficiens* USDA110-soybean symbiosis (Minder *et al.*, 2001). Previous studies also reported a possible link between phospholipids and phosphate homeostasis in rhizobia as shown in *Sinorhizobium meliloti,* where phospholipids are replaced with phosphorus-free lipids under phosphate limiting conditions (Zavaleta-Pastor *et al.*, 2010). Similarly, two phospholipases C genes are upregulated in E- and S-type bacteroids (LFC>2.2-2.9, Fig S5). Although modulation of the phosphorus level in membrane lipids does not affect nitrogen fixation in alfalfa nodules (López-Lara *et al.*, 2005), a phosphorus starvation in *Aeschynomene* nodules could be at the origin of the modulation of phospholipid metabolism in bacteroids. Indeed, the critical phosphate nutritional state of bacteroids is suggested by the high induction of phosphate and phosphonate import systems, encoded by the genes *pstSCABphoBU* (LFC=2.8-5.4) and *phnCDE1E2* (LFC=5-9.7) respectively. Interestingly, *pstB* and *pstC* genes are essential for nitrogen fixation in the *Bradyrhizobium* sp. ORS278/*A. indica* symbiosis (Bonaldi *et al.*, 2010). Furthermore, those two transporters appear to be instrumental for nitrogen fixation in *S. meliloti*, where the *pstC/phnC* (the original name of *S. meliloti phnC* was *phoC*) double mutant

is impaired for nitrogen fixation in *M. sativa* nodules (Yuan *et al.*, 2006). In bacteria, phosphonates, a family of organophosphorus compounds, can be transported in the cytoplasm by the phosphonate transporter *phnCDE1E2* and degraded into phosphate and an alkane by the carbon-phosphorus lyase complex *phnGHIJKLMN* (Hove-Jensen *et al.*, 2014; Villarreal-Chiu *et al.*, 2012). This set of genes, located next to *phnCDE1E2* is also upregulated in bacteroids (LFC=1.7-3.6). The upregulation of the *pst* and *phn* genes in the two *Aeschynomene* hosts indicates that these terminally differentiated bacteroids are phosphate starved. A recent analysis of these phosphate transport systems in different rhizobium-legume associations suggested that bacteroids are only phosphate limited when they are not terminally differentiated while terminally differentiated bacteroids do not upregulate the *pst* genes (Hu *et al.*, 2018). Our observations, however, do not confirm this proposed link between terminal bacteroid differentiation and phosphate starvation but rather suggest that the phosphate status of the bacteroids depends on other host-specific factors.

Metabolite transporters

Not surprisingly, the dicarboxylate transporter *dctA_1* gene is strongly induced (LFC>5.6), dicarboxylates being the major carbon source fueling bacteroids (Udvardi and Poole, 2013). The adjacent genes *dctBD* encoding the two-component regulator of the transporter are not differentially expressed in bacteroids. Two clusters of oligopeptide ABC transporters also displayed high (BRAD285_v2_5810-5814, LFC>6) and moderate (BRAD285_v2_6964-6967, LFC>3) induction, although their precise function remains to be determined. We hypothesized that they could be involved in the transport of NCR-like peptides as those were recently discovered in *Aeschynomene* nodule tissues (Czernic *et al.*, 2015). Two clusters of putative branched chain amino acid ABC transporters were upregulated (BRAD285_v2_3301-3305 and BRAD285_v2_0538-0542, LFC>2-5) suggesting a possible symbiotic auxotrophy of bacteroids as described in the pea/*R. leguminosarum* bv. *viciae* symbiosis (Prell *et al.*, 2009). Symbiotic auxotrophy is due to the drastic downregulation of *ilv* and *leu* genes involved in the biosynthesis of branched chain amino acids. However, *Bradyrhizobium* sp. ORS285 orthologous *ilv* and *leu* genes did not display a significant downregulation in bacteroid conditions, suggesting that the metabolic status of branched chain amino acids differs between *Bradyrhizobium* bacteroids in *Aeschynomene* and *R. leguminosarum* bacteroids in pea. Three genes potentially involved in molybdenum transport are induced in bacteroids located on the

large *nif/fix/phn* region. Those genes are *mop* coding for a Molybdenum-pterin binding protein, *modD* involved either in molybdenum transport or in signal transduction as proposed in *Rhodobacter capsulatus* (Wang *et al.*, 1993), and the putative molybdenum ABC transporter compontent encoding *modC*.

Phytosynthetic bradyrhizobia-specific genes

A large cluster of photosynthetic bradyrhizobia-specific genes is the second most induced gene cluster in bacteroids and carries 15 adjacent genes (LFC=8-11; Fig. 3B; Supporting Information Fig. S5). The cluster contains a gene tentatively annotated as *vanB* encoding one of the two subunits of vanillate O-demethylase oxidoreductase. This enzyme catalyzes the conversion of vanillate into protocatechuate and formaldehyde. These molecules are intermediates in the lignin degradation by soil microbes but are also found in plant cells and the rhizosphere (Priefert *et al.*, 1997). However, the other genes of this cluster of upregulated genes are poorly annotated. Moreover, a second *vanB* gene and an adjacent *vanA* gene which encodes the second, essential, subunit of the vanillate demethylase enzyme, are in the genome of strain ORS285 separated from the first *vanB* gene by seven other genes and are not part of the cluster of upregulated genes. Moreover, other genes of the vanillate, protocatechuate and formaldehyde degradation pathway in bradyrhizobia (Ito *et al.*, 2006) are either absent in the genome of ORS285 or are not induced in bacteroids. Therefore, the biochemical and the symbiotic function of the up-regulated gene cluster remains to be determined.

Type II secretion system

In the locus encoding the type II secretion system (T2SS), the *gspDEFGHIJKLMN* genes of the pathway are tandem duplicated. This duplication is conserved among the photosynthetic *Bradyrhizobia* but not in other *Bradyrhizobium* spp. which carry a single gene copy. This secretion system is common in gram-negative bacteria but not universal and many rhizobia including *S. meliloti* lack a T2SS.

The two copies in strain ORS285 represent each a complete T2SS machinery except that they both lack a *gspC* gene. In T2SS with a *gspC* gene, the gene was found to be essential for protein secretion. The GspC protein assures the contact between the inner membrane and outer

membrane components of the T2SS and it determines also secretion specificity (Korotkov *et al.*, 2012). Instead, the *gspD* genes of both T2SS gene copies in strain ORS285 are in a bicistronic operon with a gene encoding a secreted Sel1 domain protein which is conserved in *gspD* operons in other bacteria lacking *gspC* suggesting that the Sel1 domain protein could replace the function of GspC. The *gspS* gene encoding an outermembrane lipoprotein called pilotin is also missing in both copies of strain ORS285 but this is not uncommon among T2SS (Korotkov *et al.*, 2012). Finally, the locus carries only one copy of the gene encoding the essential prepilin signal peptidase.

Both T2SS encoding copies of strain ORS285 are induced in bacteroids but the first copy is expressed at low levels while the second copy is expressed to up to more than hundred-fold higher levels in bacteroids (Supporting Information Fig. S6). T2SS secretes proteins to the extracellular environment. Substrates are very often enzymes and depending on the bacterial species the number of substrates requiring the T2SS range from one protein to several tens. The T2SS locus of strain ORS285 carries in addition to the core components of the apparatus, a few genes that potentially encode secretion substrates. These genes are also strongly or very strongly induced in bacteroids (LFC=6-9). They encode proteins that are only vaguely annotated. One of them is a phosphoesterase and a second one is a metallo-dependent phosphatase suggesting that these two proteins are involved in the dephosphorylation of a further not determined substrate. Another potential T2SS substrate protein encoded by BRAD285_v2_1078 is particularly interesting because it is among the most highly induced (LFC=9.2 in *A. afraspera* or 6.2 in *A. indica*) genes in bacteroids (34th most expressed gene, among the first percentile genes) in *A. afraspera* suggesting its importance. The protein is well conserved among bradyrhizobia and some other gram-negative bacteria, but its homologs have no known function. However, the protein has also a weaker homology of 27-34% with S-layer proteins of some gram-positive bacteria. S-layers are proteinaceous surface layers, which are formed by the self-assembly of monomeric proteins into a regularly spaced, two-dimensional array. S-layer proteins are highly diverse between bacterial species and are found in both Gram-positive and Gram-negative bacteria (Fagan and Fairweather, 2014). In Gram-negative bacteria, S-layers are attached to the LPS and, depending on the bacterial species, the S-layer proteins can be secreted by a dedicated T2SS. This suggests thus that the BRAD285_v2_1078-encoded protein forms a bacteroid specific S-layer. Its very high expression would be compatible with such a function. Among the functions of S-layers is the interaction with the host immune system. For example, in *Caulobacter crescentus*, an alphaproteobacterium closely related to the

rhizobia, the S-layer protects against antimicrobial peptides (de la Fuente-Núñez *et al.*, 2012). It is easy to imagine that a putative BRAD285_v2_1078-determined S-layer could provide protection against the NCR peptides that target the bacteroids in *Aeschynomene* nodules (Czernic *et al.*, 2015).

The T2SS substrate proteins are secretory proteins that must first be translocated from the cytosol to the periplasm via either the SEC translocon or via the twin-arginine protein translocon (TAT). The genes encoding these two different translocons in *Bradyrhizobium* strain ORS285 are not differentially expressed in the bacteroids (Supporting Information Fig. S6).

Symbiosis island

Bradyrhizobium strain ORS285, has a clearly defined island carrying the nodulation genes and a T3SS, suggesting that it evolved from a nitrogen fixing strain interacting with the Nod factor independent *Aeschynomene* species and that the acquisition of symbiosis island carrying nodulation genes through horizontal gene transfer had been at the origin of a host range extension to Nod factor-dependent *Aeschynomene* plants. The symbiotic island has been inserted in a Met-tRNA gene, resulting in a 50 bp direct repeat flanking the island, a typical configuration for genomic islands. The island contains 374 predicted open reading frames from which 204 are derived from transposons and another 116 have an unknown function, many of which might also be related to insertion sequences. Thus, the larger part of the symbiotic island consists of transposons, another typical feature of genomic islands. The remaining 54 genes of the island are mainly composed of the nodulation genes (20 genes) and a type III secretion system (T3SS) (22 genes).

Cell cycle and peptidoglycan synthesis

Peptidoglycan, and in particular the regulation of peptidoglycan-remodeling enzymes, is the primary cell-shape determinant of bacteria (Typas *et al.*, 2012; van Teeseling *et al.*, 2017). Thus, the morphological changes in E- and S-type bacteroids might be accompanied with changes in peptidoglycan synthesis. Accordingly, we recently revealed the essential role of a DD-carboxypeptidase 1 (DD-CPase 1), a murein hydrolase, in symbiosis. The absence of this protein leads to hypertrophy and death of *Bradyrhizobium* bacteroids in *Aeschynomene* nodules

(Gully *et al.*, 2016; Barrière *et al.*, 2017). On the other hand, the cell cycle of *Alphaproteobacteria* is a tightly regulated process in which genome replications are alternated by cell divisions; the polyploidization implies that the usual cell cycle regulation has been abandoned and replaced by a new regulation mode. We reasoned that these processes should be reflected in the transcriptome of bacteroids. Therefore, we used the literature on *Alphaproteobacteria* (Biondi *et al.*, 2006; Brilli *et al.*, 2010; Cameron *et al.*, 2014; de Nisco *et al.*, 2014) to assemble most of the peptidoglycan biosynthesis and regulation and the cell cycle regulatory genes of strain ORS285 and analyzed their expression in our transcriptome dataset.

Concerning cell division regulatory genes, overall about half of the considered genes (22 out of 48) show a differential (up or down) expression in bacteroids compared to free-living bacteria, confirming the cell cycle switch towards endoreduplication in both elongated and spherical terminally differentiated bacteroids (Supporting Information Fig. S8). The master cell cycle regulatory gene *ctrA* is downregulated in bacteroid conditions (|LFC|>2), as observed in *S. meliloti* bacteroids or after NCR challenge (Penterman *et al.*, 2014; Roux *et al.*, 2014; Pini *et al.*, 2015). CtrA represses DNA replication and promotes cell division (Biondi *et al.*, 2006; Pini *et al.*, 2015). In *C. crescentus*, CtrA controls the expression of *ftsZ*, *ftsA* and *ftsQ* genes (Laub *et al.*, 2002), which mediate septum formation (Haeusser and Margolin, 2016). In *S. meliloti*, these CtrA targets are downregulated in differentiated bacteroids. However, while those genes are differentially repressed in S-type bacteroids, *ftsZ* (|LFC|=1.20) and *ftsQ* (|LFC|=1.49) are below the DEG cutoff in E-type bacteroids. This may be due to the existence of a mixed population of differentiated bacteroids and undifferentiated bacteria in *A. afraspera* nodules (Bonaldi *et al.*, 2011). On the other hand, we observed an upregulation of the *minCDE* operon (LFC>1.64), which inhibit FtsZ assembly into a septal ring (Rowlett and Margolin, 2015). Taken together, these results highlight the importance of modulating the cell cycle regulation to obtain the DNA polyploidization and cell division inhibition in E- and S-type bacteroids.

A shift in peptidoglycan biosynthesis was also visible in the transcriptome although less pronounced than the cell cycle switch (Supporting Information Fig. S9). Among the remarkable genes are *lrgAB* which are strongly upregulated in bacteroids (LFC>3.6). These genes are widely conserved but only functionally characterized in *Staphylococcus* and *Streptococcus* spp. as inhibitors of murein hydrolase activity (Groicher *et al.*, 2000). The *murA* gene encoding the first enzyme of peptidoglycan precursor synthesis is also induced in bacteroids (LFC>2.3). Finally, another interesting class of differentially expressed genes are encoding L,D-transpeptidase (LD-TPase) proteins, enzymes involved in the formation of peptide bridges

between peptidoglycan strands. *Rhizobiales* contain an unusual abundance of putative LD-TPase-encoding genes (Cameron *et al.*, 2014). We identified twelve of them in the genome of stain ORS285. Two of the LD-TPase genes are repressed in bacteroids while four of them are upregulated (Supporting Information Fig. S9). Remarkably, the DD-CPase 1 gene, genetically identified as essential for bacteroid differentiation, is not differentially regulated in bacteroids. Altogether, this genetic and transcriptome evidence highlight the importance of the cell wall dynamics in the cell enlargement during bacteroid differentiation.

Bacterial envelope

Among previously identified functions of photosynthetic *Bradyrhizobia* involved in the chronic infection of *Aeschynomene* nodules are the lipopolysaccharide (LPS) and hopanoid lipid components of the cell envelope (Silipo *et al.*, 2014; Kulkarni *et al.*, 2015; Busset *et al.*, 2016; Busset *et al.*, 2017). Mutants that are affected in the production of these lipids form defective bacteroids. Therefore, it was of interest to specifically check the expression pattern of genes involved in the biosynthesis of these membrane molecules. Of the ten hopanoid biosynthesis genes identified, only three, *hpnO*, *hpnP* and *hpnG* are slightly but significantly upregulated in bacteroids (LFC=1.7-2.4) (Supporting Information Fig. S10). These genes encode enzymes that modify the hopanoid skeleton, introducing amine, methyl and glycosyl groups. Such modifications affect the biophysical properties of the lipids and the membranes in which they are inserted (Belin *et al.*, 2018). Thus, the upregulation of these three genes suggests that the bacteroids use hopanoids to adjust the properties (rigidity) of their membranes in accordance to the local stress condition.

Analysis of the LPS biosynthesis genes (Supporting Information Fig. S11) does not suggest any notable LPS difference between free-growing bacteria and bacteroids except for the *lpxXL* gene which is slightly repressed in bacteroids (LFC=-1,6/-1,9) even if it maintains a high expression. This is surprising because this gene, which encodes an acyltransferase that introduces one of the very long-chain fatty acids in the lipid A moiety of LPS, is required for normal bacteroid formation (Busset *et al.*, 2017).

In addition, the expression profiles were analyzed for all other genes with annotations related to the biosynthesis of uncharacterized polysaccharides (*e.g.* genes encoding glycosyl transferases and nucleotide sugar biosynthesis enzymes). This analysis revealed a few enzymes

with altered expression, either repressed or induced (Supporting Information Fig. S12) but the nature of the affected polysaccharides and the low expression levels of these genes does not allow to evaluate the significance of these differences.

Downregulated genes in bacteroids

Several sets of genes were downregulated in bacteroids. Striking examples are genes encoding the functions for pilus assembly (two loci, one carrying the *cpaABCDEF* genes and the other carrying a second copy of the *cpaBCDE* genes, |LFC|=2.4-6.4). Also, flagellar motility and chemotaxis were, not surprisingly, strongly downregulated in bacteroids (Supporting Information Fig. S13). Downregulation of the genes encoding ribosomal proteins of the 30S and 50S subunits (|LFC|=1-3.4; Supporting Information Fig. S14) indicates that protein translation is slowed down in bacteroids, a finding that has been reported before in bacteroids of other rhizobia and that suggests that the household metabolism is slowed down in bacteroids (Pessi *et al.*, 2007; Karunakaran *et al.*, 2009). This downregulation could be linked to plant NCR secretion. Indeed, NCR247 and NCR335 are responsible to the decreased expression of ribosomal proteins in *S. meliloti* 1021 (Tiricz *et al.*, 2013). Unexpectedly, downregulation of bacterial photosynthesis genes, including the *cbb* genes of the Calvin-Benson-Bassham cycle, the *puf* genes encoding the photosynthetic reaction center and the *bch* genes for bacteriochlorophyll biosynthesis, was also observed (|LFC|=1.58-6.9) (Supporting Information Fig. S15), despite the fact that the plant roots and nodules were exposed to light in the transparent tubes used for plant growth. This downregulation also contrasts with the positive effect of *Bradyrhizobium* sp. ORS278 photosynthesis on stem nodulation of *Aeschynomene sensitiva* (Giraud *et al.*, 2000), and the requirement of one of the two bacterial RubisCO genes for the establishment of a functional symbiosis with *A. indica* (Gourion *et al.*, 2011). In nitrogen-fixing symbiosis, RubisCO and by extension the Calvin–Benson–Bassham cycle, were proposed to play a role in the formation of a transient electron sink until the nitrogenase is becoming active. Indeed, a high level of redox power is transferred to the bacteroids via the dicarboxylate import and integration in the TCA cycle, and RubisCO could balance a critical redox state before the establishment of a functional nitrogenase activity (Gourion *et al.*, 2011). Our transcriptome data, together with the previous genetic analysis of the CO_2 fixation pathway indicate that this process is only essential at the early stages of the interaction and that once symbiosis is fully activated, CO_2 fixation is not essential anymore.

Strain ORS285 has the capacity to produce several secondary metabolites. Analysis with the antiSMASH algorithm revealed the presence of one N-acyl homoserine lactone (AHL) synthesis cluster producing a quorum sensing signal and four non-ribosomal peptide synthetase (NRPS) gene clusters determining the production of so far uncharacterized peptides. The AHL cluster as well as 3 NRPS clusters out of 4 are downregulated under bacteroid conditions (Supporting Information Fig. S16).

References

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**: 9063.

Belin, B.J., Busset, N., Giraud, E., Molinaro, A., Silipo, A., and Newman, DK. (2018) Hopanoid lipids: from membranes to plant–bacteria interactions. *Nat Rev Microbiol* **16**: 304–315.

Biondi, E.G., Reisinger, S.J., Skerker, J.M., Arif, M., Perchuk, B.S., Ryan, K.R., *et al.* (2006) Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature* **444**: 899–904.

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.

Bonaldi, K., Gourion, B., Fardoux, J., Hannibal, L., Cartieaux, F., Boursot, M., *et al.* (2010) Large-scale transposon mutagenesis of photosynthetic *Bradyrhizobium* sp. strain ORS278 reveals new genetic loci putatively important for Nod-independent symbiosis with *Aeschynomene indica*. *Mol Plant Microbe Interact* **23**: 760–770.

Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., *et al.* (2010) The diversity and evolution of cell cycle regulation in alpha-proteobacteria: A comparative genomic analysis. *BMC Syst Biol* **4**: 52.

Busset, N., De Felice, A., Chaintreuil, C., Gully, D., Fardoux, J., Romdhane, S., *et al.* (2016) The LPS O-antigen in photosynthetic *Bradyrhizobium* strains is dispensable for the

establishment of a successful symbiosis with *Aeschynomene* legumes. *PLoS ONE* **11**: e0148884.

Busset, N., Di Lorenzo, F., Palmigiano, A., Sturiale, L., Gressent, F., Fardoux, J., *et al.* (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. *Front Microbiol* **8**: 1821.

Cameron, T.A., Anderson-Furgeson, J., Zupan, J.R., Zik, J.J., and Zambryski, P.C. (2014) Peptidoglycan synthesis machinery in *Agrobacterium tumefaciens* during unipolar growth and cell division. *mBio* **5**: e01219-14.

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, *et al.* (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.

de la Fuente-Nunez, C., Mertens, J., Smit, J., and Hancock R.E.W. (2012) The Bacterial surface layer provides protection against antimicrobial peptides. *Appl Environ Microbiol* **78**: 5452–5456.

De Nisco, N.J., Abo, R.P., Wu, C.M., Penterman, J., and Walker, G.C. (2014) Global analysis of cell cycle gene expression of the legume symbiont *Sinorhizobium meliloti*. *Proc Natl Acad Sci U S A* **111**: 3217–3224.

Fagan, R.P., and Fairweather, N.F. (2014) Biogenesis and functions of bacterial S-layers. *Nat Rev Microbiol* **12**: 211–222.

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.

Gourion, B., Delmotte, N., Bonaldi, K., Nouwen, N., Vorholt, J.A., and Giraud, E. 2011. Bacterial RuBisCO is required for efficient *Bradyrhizobium*/*Aeschynomene* symbiosis. *PLoS ONE* **6**: e21900.

Groicher, K.H., Firek, B.A., Fujimoto, D.F., and Bayles, K.W. (2000) The *Staphylococcus aureus lrgAB* operon modulates murein hydrolase activity and penicillin tolerance. *J Bacteriol* **182**: 1794–1801.

Gully, D., Gargani, D., Bonaldi, K., Grangeteau, C., Chaintreuil, C., Fardoux, J. *et al.* (2016) A peptidoglycan-remodeling enzyme is critical for bacteroid differentiation in *Bradyrhizobium* during legume symbiosis. *Mol Plant Microbe Interact* **29**: 447–457.

Hove-Jensen, B., Zechel, DL., and Jochimsen, B. (2014) Utilization of glyphosate as phosphate source: biochemistry and genetics of bacterial carbon-phosphorus lyase. *Microbiol Mol Biol Rev* **78**: 176–197.

Hu, Y., Jiao, J., Liu, L.X., Sun, Y.W., Chen, W.F., Sui X.H., *et al.* (2018) Evidence for phosphate starvation of rhizobia without terminal differentiation in legume nodules. *Mol Plant Microbe Interact* doi: 10.1094/MPMI-02-18-0031-R.

Ito, N., Itakura, M., Eda, S., Saeki, K., Oomori, H., Yokoyama, T., *et al.* (2006) Global gene expression in *Bradyrhizobium japonicum* cultured with vanillin., vanillate., 4-hydroxybenzoate and protocatechuate. *Microbes Environ* **21**: 240–250.

Karunakaran, R., Ramachandran, V.K., Seaman, JC., East, AK., Mouhsine, B., Mauchline, T.H., *et al.* (2009) Transcriptomic analysis of *Rhizobium leguminosarum* biovar viciae in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. *J Bacteriol* **191**: 4002–4014.

Korotkov, K.V., Sandkvist, M., and Hol, W.G.J. (2012) The type II secretion system: Biogenesis., molecular architecture and mechanism. *Nature Reviews Microbiology* **10**: 336–351.

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 B*radyrhizobium diazoefficiens*. *mBio* **6**: e01251-15.

Laub, M.T., Chen, S.L., Shapiro, L., and McAdams, H.H. (2002) Genes directly controlled by CtrA., a master regulator of the *Caulobacter* cell cycle. *Proc Natl Acad Sci U S A* **99**: 4632–4637.

López-Lara, I.M., Gao, J-L., Soto, M.J., Solares-Pérez, A., Weissenmayer, B., Sohlenkamp, C., *et al.* (2005) Phosphorus-free membrane lipids of *Sinorhizobium meliloti* are not required for the symbiosis with alfalfa but contribute to increased cell yields under phosphorus-limiting conditions of growth. *Mol Plant Microbe Interact* **18**: 973–982.

Minder, A.C., De Rudder, K.E.E., Narberhaus, F., Fischer, H-M., Hennecke, H., and Geiger, O. (2001) Phosphatidylcholine levels in *Bradyrhizobium japonicum* membranes are critical for an efficient symbiosis with the soybean host plant. *Mol Microbiol* **39**: 1186–1198.

Penterman, J., Abo, R.P., De Nisco, N.J., Arnold, M.F.F., Longhi, R., Zanda, M., *et al.* (2014) Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. *Proc Natl Acad Sci U S A* **111**: 3561–3566.

Pessi, G., Ahrens, CH., Rehrauer, H., Lindemann, A., Hauser, F., Fischer, H-M., *et al.* (2007) Genome-wide transcript analysis of *Bradyrhizobium japonicum* bacteroids in soybean root nodules. *Mol Plant Microbe Interact* **20**: 1353–1363.

Pini, F., De Nisco, N.J., Ferri, L., Penterman, J., Fioravanti, A., Brilli, M., *et al.* (2015) Cell cycle control by the master regulator CtrA in *Sinorhizobium meliloti*. *PLoS Genet* **11**: 1–24.

Prell, J., White, J.P., Bourdes, A., Bunnewell, S., Bongaerts, R.J., and Poole P.S. (2009) Legumes regulate Rhizobium bacteroid development and persistence by the supply of branched-chain amino acids. *Proc Natl Acad Sci U S A* **106**: 12477–12482.

Priefert, H., Rabenhorst, J., Steinbuchel, A., and Rabenhorst, R. (1997) Molecular characterization of genes of *Pseudomonas* sp . strain HR199 involved in bioconversion of vanillin to protocatechuate. *J Bacteriol* **179**: 2595–2607.

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**: 817–837.

Rowlett, V.W., and Margolin, W. (2015) The Min system and other nucleoid-independent regulators of Z ring positioning. *Frontiers in Microbiology* **6**: 478.

Silipo, A., Vitiello, G., Gully, D., Sturiale, L., Chaintreuil, C., Fardoux, J., *et al.* (2014) Covalently linked hopanoid-lipid A improves outer-membrane resistance of a Bradyrhizobium symbiont of legumes. *Nat Commun* **5**: 5106.

van Teeseling, M.C.F., de Pedro, M.A., and Cava, F. (2017) Determinants of bacterial morphology: From fundamentals to possibilities for antimicrobial targeting. *Front Microbiol* **8**: 1264.

Tiricz, H., Szücs, 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.

Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2011) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol* **10**: 123–136.

Udvardi, M., and Poole, P.S. (2013) Transport and metabolism in legume-rhizobia symbioses. *Annu Rev Plant Biol* **64**: 781–805.

Villarreal-Chiu, J.F., Quinn, J.P., and McGrath, J.W. (2012) The genes and enzymes of phosphonate metabolism by bacteria., and their distribution in the marine environment. *Front Microbiol* **3**: 19.

Wang, G., Angermuller, S., and Klipp, W. (1993) Characterization of *Rhodobacter capsulatus* genes encoding a molybdenum transport system and putative molybdenum-pterin-binding proteins. *J Bacteriol* **175**: 3031–3042.

Yuan, Z., Zaheer, R., and Finan, T.M. (2006) Regulation and properties of PstSCAB, a transport system of *Sinorhizobium meliloti*. *J Bacteriol* **188**: 1089–1102.

Zavaleta-Pastor, M., Sohlenkamp, C., Gao, J-L., Guan, Z., Zaheer, R., Finan, T.M., *et al.* (2010) *Sinorhizobium meliloti* phospholipase C required for lipid remodeling during phosphorus limitation. *Proc Natl Acad Sci U S A* **107**: 302–307.

Appendix S3 Methods for transcriptome analysis.

RNA extraction, transcriptomics and qRT-PCR

Nodules were harvested at 14 dpi, frozen in liquid N_2 and kept at -80°C until RNA extraction. For transcriptomic analysis, total nodule RNA was extracted as previously described (Mergaert *et al.*, 2003), and submitted to plant and bacterial rRNA and eukaryotic mRNA depletion using RiboZero kits for plants and for Gram-negative bacteria (Epicentre, Madison, USA). Total RNA of free-living cells was extracted from bacterial pellets obtained from cultures grown in YM medium at OD_{600}=0.5 using the hot phenol and hot SDS method (Chapelle *et al.*, 2015). Oriented libraries of bacterial mRNA were produced using the SOLiD total RNA-seq kit and sequenced on a SOLiD 5500XL (50bp per read). Each condition was

performed in biological triplicates, resulting in the sequencing of 9 libraries which yielded ca. 30 million reads per sample, out of which ca. 9 million per sample were unique bacterial gene reads. The sequencing data of this project was submitted to the Sequence Read Archive (SRA) and obtained accession n° GSE108744 (SRA number: SRP128034).

For qRT-PCR analyses, total RNA was extracted using the Plant RNeasy kit (Qiagen). DNase treatment was performed on 2 µg RNA using the TurboDNA-free kit (Ambion, Life Technologies). Reverse transcription was performed with a RevertAid Reverse Transcriptase kit (Fermentas, Thermo Scientific). qRT-PCRs were carried out on a Roche LightCycler® 96 System using the Light Cycler Fast Start DNA Master SYBR Green I kit (Roche, http://www.roche.com/) and the primers that are listed in Supporting Information Table S1. All kits were used according to the manufacturer protocols. Cycling conditions were as follows: 95°C for 10 min, 50 cycles at 95°C for 5 sec, 58°C for 5 sec, and 72°C for 15 sec. The highly expressed 16S rDNA gene (*rrs*) and the moderately expressed BRAD285_v2_7630 gene annotated as *miaB* have been taken as references for qRT-PCR analyses. They have been chosen due to their invariant expression among conditions according to our transcriptome dataset, and their contrasted expression levels.

Statistical and differential expression analyses

CLC Workbench 10 software (CLC bio, Aarhus, Denmark) was used for trimming, sizing, and mapping the reads on the reference genome of *Bradyrhizobium* sp. ORS285 strain (Gully *et al.*, 2017). Only unique gene reads mapping on 80% of their length and displaying a 90% identity have been conserved for differential analyses. Count tables have then been filtered to retain only genes with a gene count over 1 count per million (cpm) in half of the samples of the dataset. Normalization and differential analyses were performed using generalized linear models as described in the DESeq2 package (version 1.12.4, Love *et al.*, 2014). Comparisons of gene expression values were performed using Wald-tests from DESeq2 package. The cutoff chosen for differentially expressed genes (DEG) are a False Discovery Rate (FDR) < 0.01 and a fold-change > 3 ($\log_2$ fold change (LFC) > 1.58). Heatmaps were produced using the ComplexHeatMap package (v 1.12, Gu *et al.*, 2016). Read repartition, based on Cluster of Orthologous Genes (COG) database, has been calculated as follow: $\text{COG class}_i = \sum \text{reads COG}_i / \sum \text{reads}$.

Genome mining analyses

An antiSMASH query was performed on the *Bradyrhizobium* sp. ORS285 genome to identify possible enzymes involved in secondary metabolite synthesis, using the MiscroScope platform (Weber *et al.*, 2015; Vallenet *et al.*, 2017).

References

Chapelle, E., Alunni, B., Malfatti, P., Solier, L., Pédron, J., Kraepiel, Y., *et al.* (2015) A straightforward and reliable method for bacterial in planta transcriptomics: Application to the *Dickeya dadantii*/*Arabidopsis thaliana* pathosystem. *Plant J* **82**: 352–362.

Gu, Z., Eils, R., and Schlesner, M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**: 2847–2849.

Gully, D., Teulet, A., Busset, N., Nouwen, N., Fardoux, J., Rouy, Z., *et al.* (2017) Complete genome sequence of *Bradyrhizobium* sp. ORS285., a photosynthetic strain able to establish Nod factor-dependent or Nod factor-independent symbiosis with *Aeschynomene* legumes. *Genome Announc* **5**: e00421-17.

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

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**: 161–173.

Vallenet, D., Calteau, A., Cruveiller, S., Gachet, M., Lajus, A., Josso, A., *et al.* (2017) MicroScope in 2017: An expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res* **45**: D517–D528.

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Bruccoleri, R., *et al.* (2015) AntiSMASH 3.0-A comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* **43**: W237–W243.

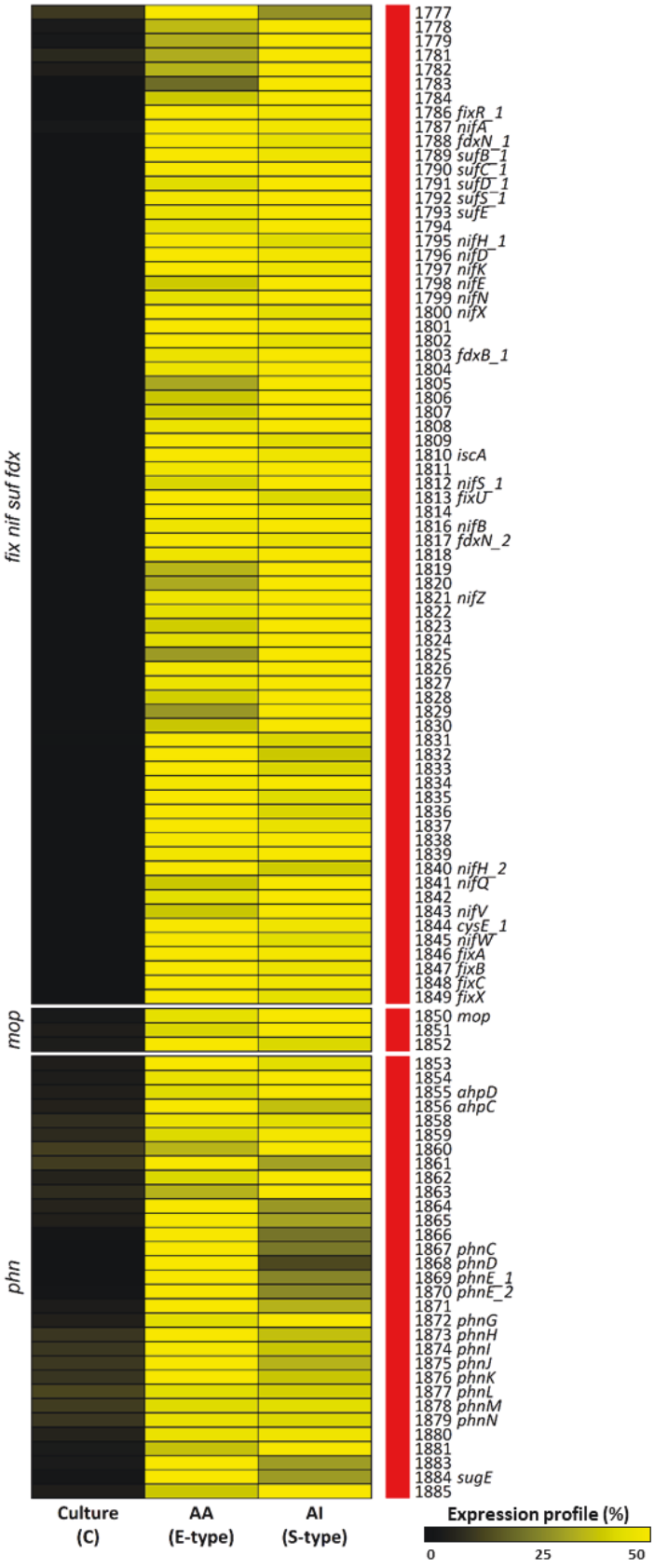


Fig S1. Nitrogenase region of *Bradyrhizobium* sp. ORS285 is highly induced in bacteroids. Heatmap showing the gene expression of the region containing *ca.* 100 genes involved in the nitrogen fixation process (LFC from 3 to 15). These genes, the *nif*, *fix* and *suf* genes, encode the nitrogenase structural subunits, chaperones for nitrogenase biogenesis and enzymes for the biosynthesis of the nitrogenase Fe-S and Fe-S-molybdenum cofactors. Together, the transcripts of this gene cluster represent about 22% of all transcripts in bacteroids (excluding ribosomal RNA and tRNA transcripts) and about half of these are allocated to the *nifH1H2DK* genes encoding the nitrogenase enzyme itself. Directly adjacent to those genes are the *mop* genes, involved in molybdenum binding, and the *phn* cluster, involved in phosphonate transport and metabolism. The right column indicates whether the gene is DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

Fig S2. The *Bradyrhizobium* sp. ORS285 locus containing *pst* and *hup* genes is induced in bacteroid conditions. Heatmap showing the gene expression of the *pst*, *hup* and *hyp* genes. The *pst* operon, involved in phosphate import and metabolism is induced in bacteroid conditions, and slightly upregulated in E-type bacteroids. *hup* genes encodes for an uptake hydrogenase and accessory functions and are upregulated in bacteroids and slightly more in S-type bacteroids. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

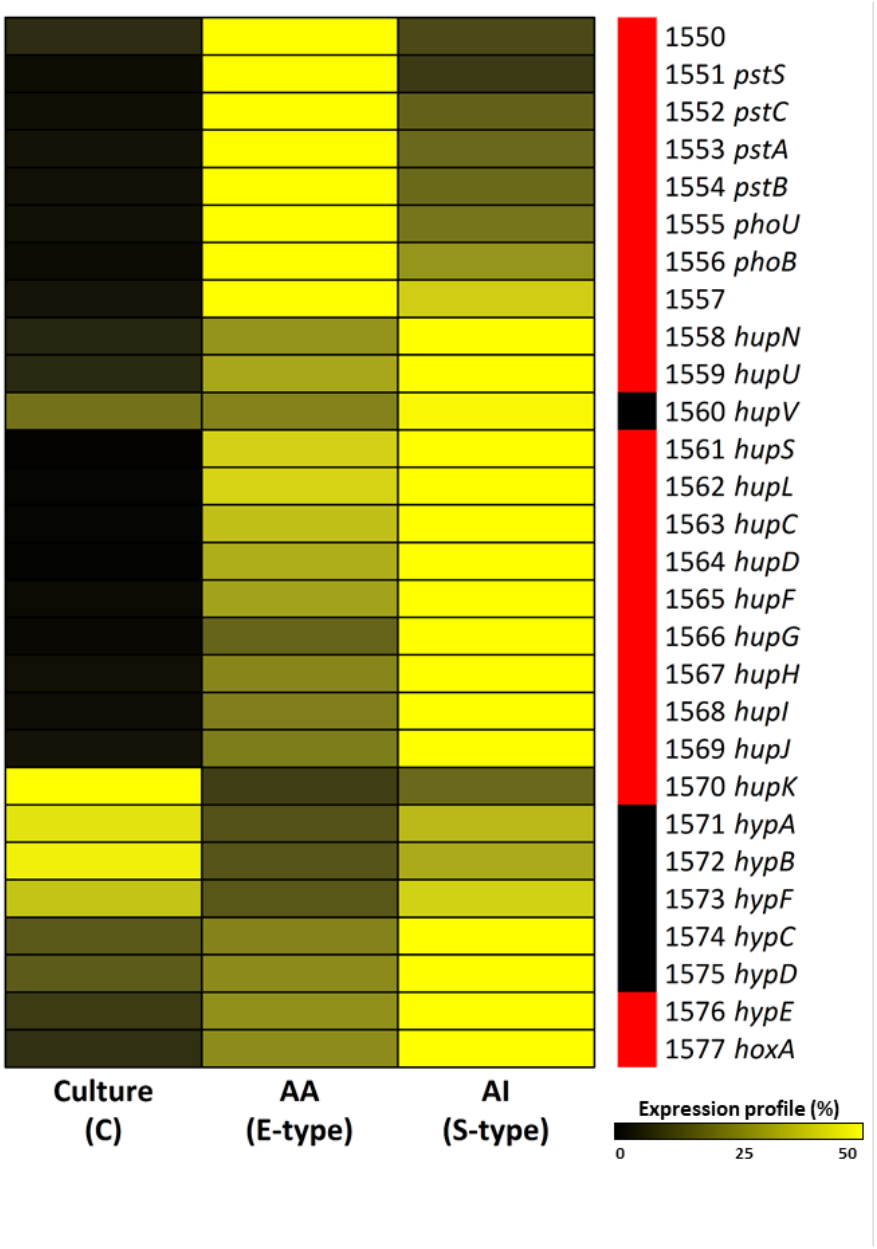


Fig S3. *Bradyrhizobium* sp. ORS285 bacteroids highly express a region involved in phospholipid metabolism as well as phospholipases C. Heatmap showing the gene expression of the *pmtA-like* cluster and two phospholipases C (*plcN*). The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

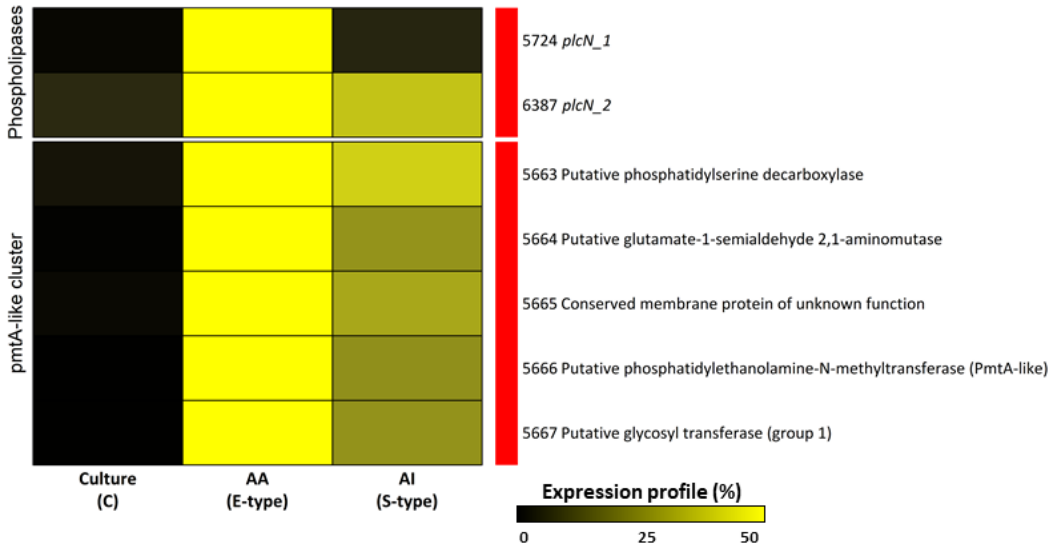


Fig S4. Expression of photosynthetic *Bradyrhizobium* specific genes. Heatmap showing the expression of the *Bradyrhizobium* sp. ORS285 genes present in *Aeschynomene* symbionts establishing a nod-independent symbiosis and absent in other bradyrhizobia. Genes are shared by *Bradyrhizobium* sp. ORS285, *Bradyrhizobium* sp. ORS278, *Bradyrhizobium* sp. BTAi1, *Bradyrhizobium oligotrophicum* S58, *Bradyrhizobium* sp. ORS375 and *Bradyrhizobium* sp. STM3809. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information

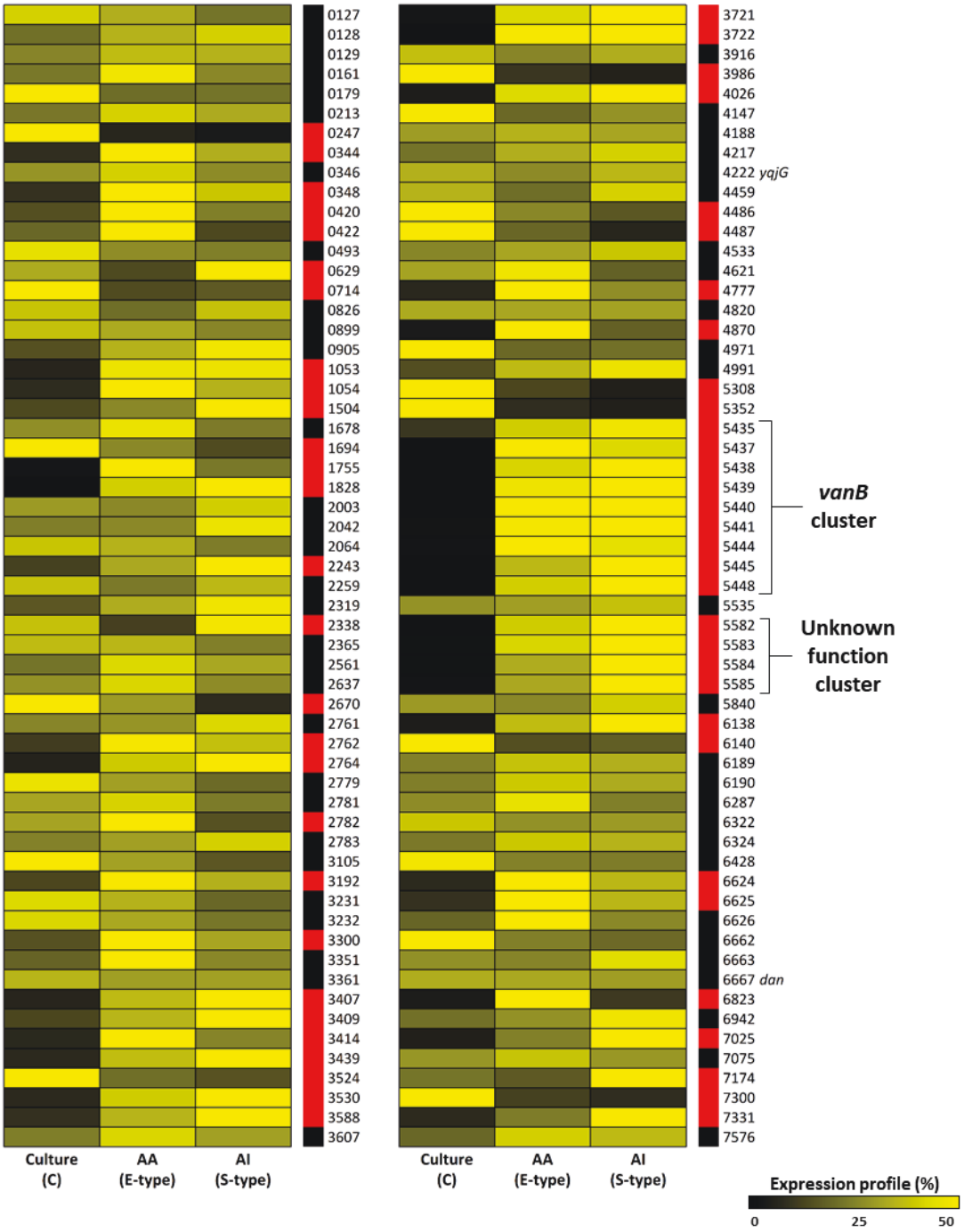


Table S2).

Fig S5. Expression of the photosynthetic *Bradyrhizobium*-specific *vanB* gene cluster in *Bradyrhizobium* sp. ORS285. Heatmap showing the gene expression of the *vanB* cluster. This cluster is the second most induced one after the genes required for nitrogen fixation in both E-type and S-type bacteroids. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

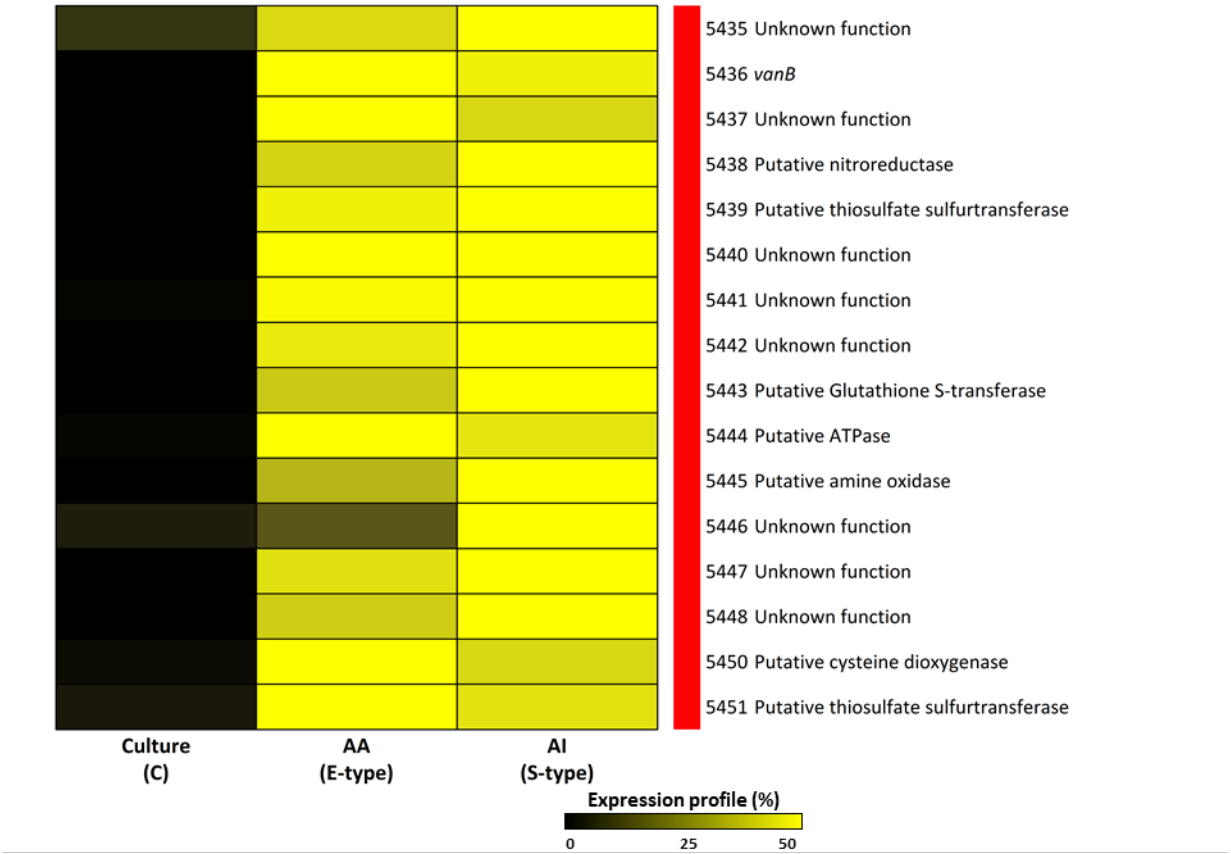


Fig S6. *Bradyrhizobium* sp. ORS285 features a duplicated Type II Secretion System (T2SS) induced in bacteroids. Heatmap showing the gene expression of genes involved in the general secretion pathway (*gsp*) or T2SS, as well as *tat* and *sec* genes. While both *gsp* clusters are induced in bacteroids, the *gsp_1* cluster expression level is ten times lower than that of the *gsp_2* cluster. The latter cluster, as well as the adjacent genes around it (from BRAD285_v2_1057 to BRAD285_v2_1080) are upregulated in E-type bacteroids. Annotation in cyan displays the two duplicated *gsp* clusters. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

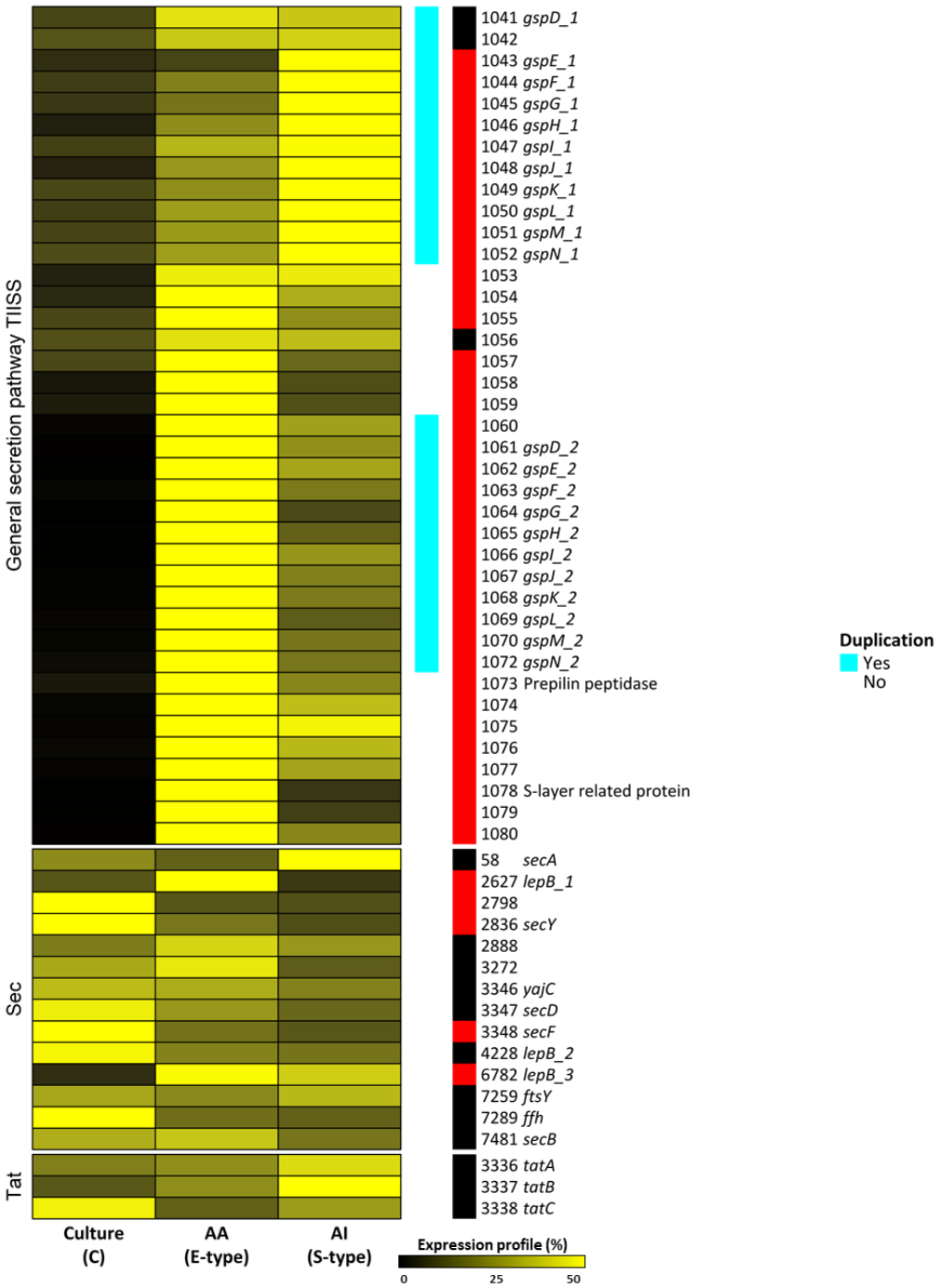


Fig S7. Expression of the *Bradyrhizobium* sp. ORS285 symbiotic island. Heatmap showing the gene expression of the symbiotic island. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

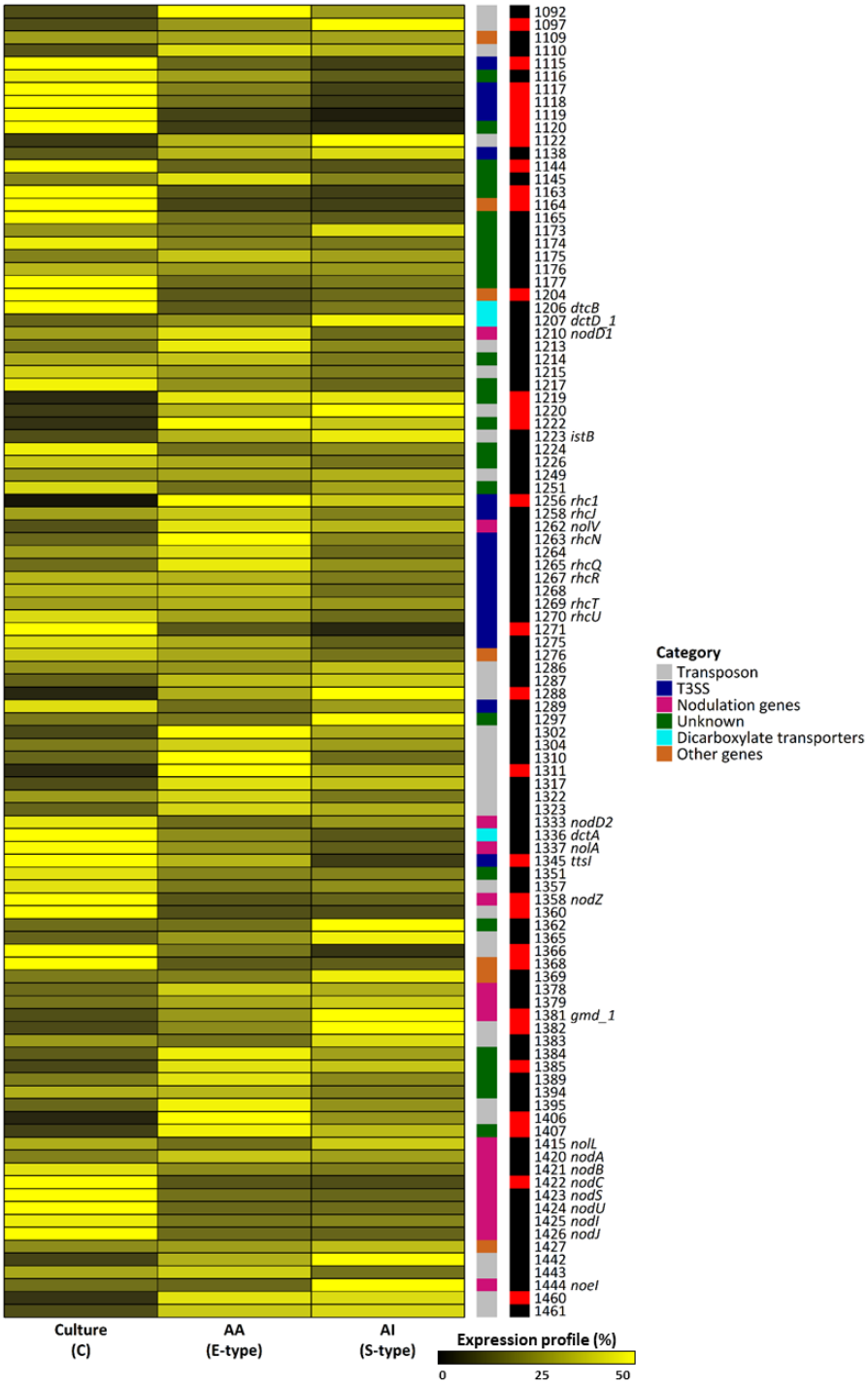


Fig S8. Expression of cell cycle regulatory genes. Heatmap showing the expression of genes involved in cell cycle control. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

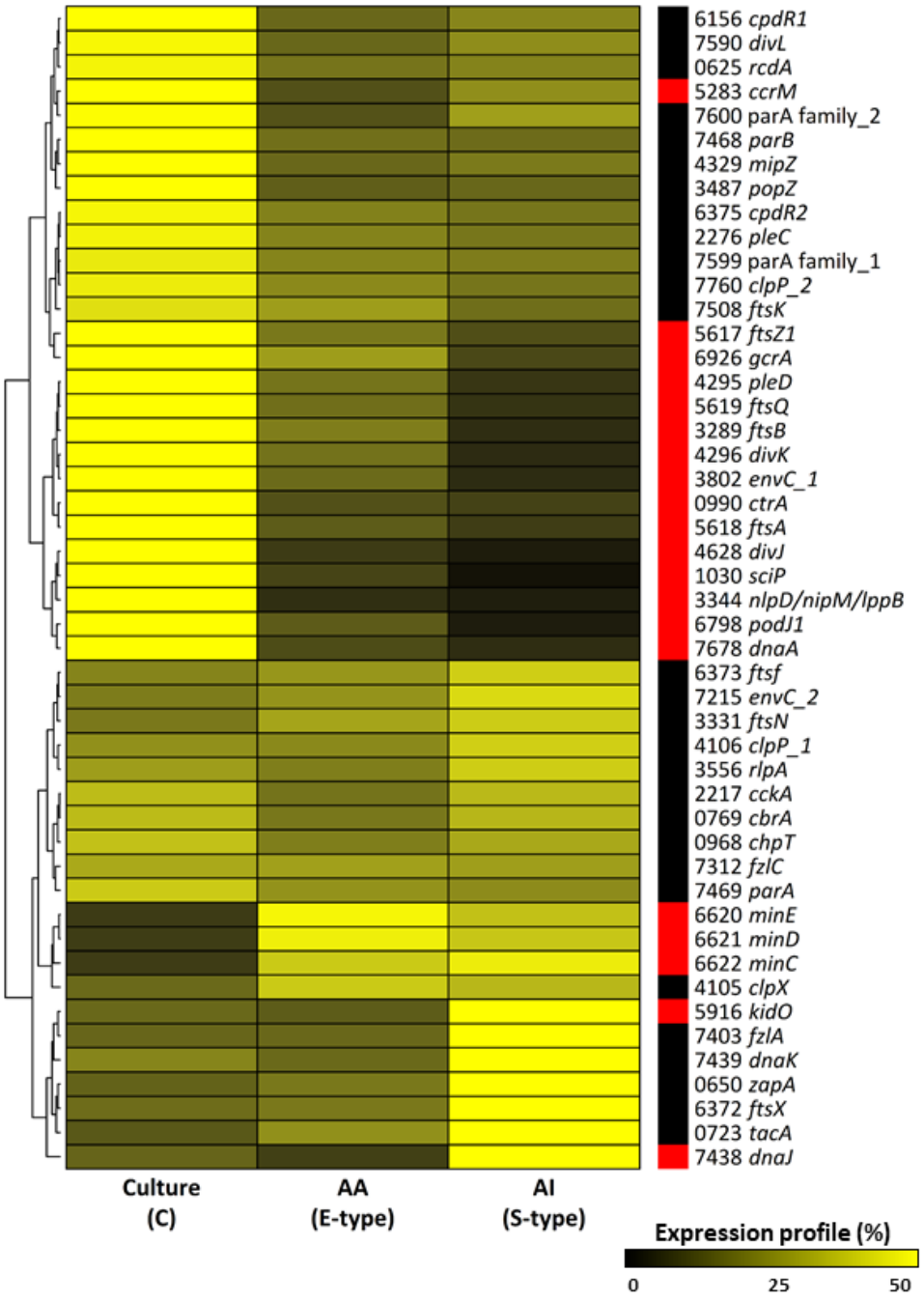


Fig S9. Expression of peptidoglycan biosynthesis genes. Heatmap showing the expression of genes involved in peptidoglycan synthesis. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

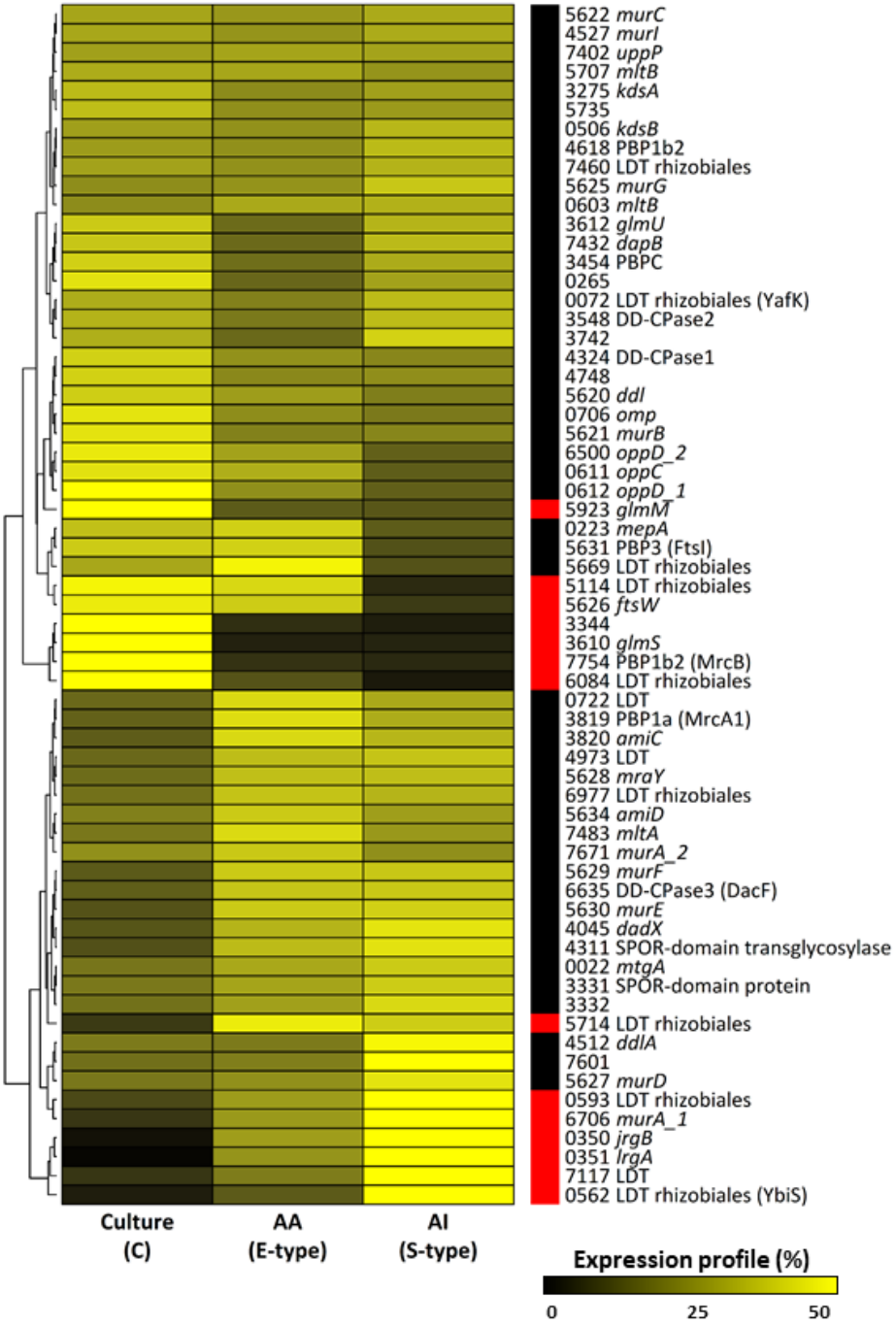


Fig S10. Expression of the hopanoid biosynthesis pathway. Heatmap showing the expression of genes involved in hopanoid synthesis. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

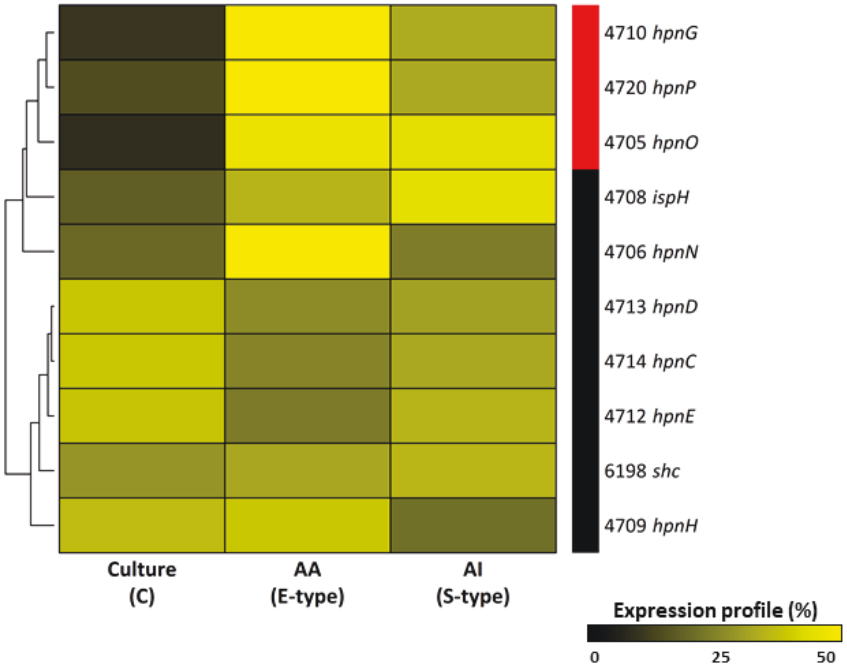


Fig S11. Expression of the lipopolysaccharide (LPS) biosynthesis pathway. Heatmap showing the expression of genes involved in LPS synthesis. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

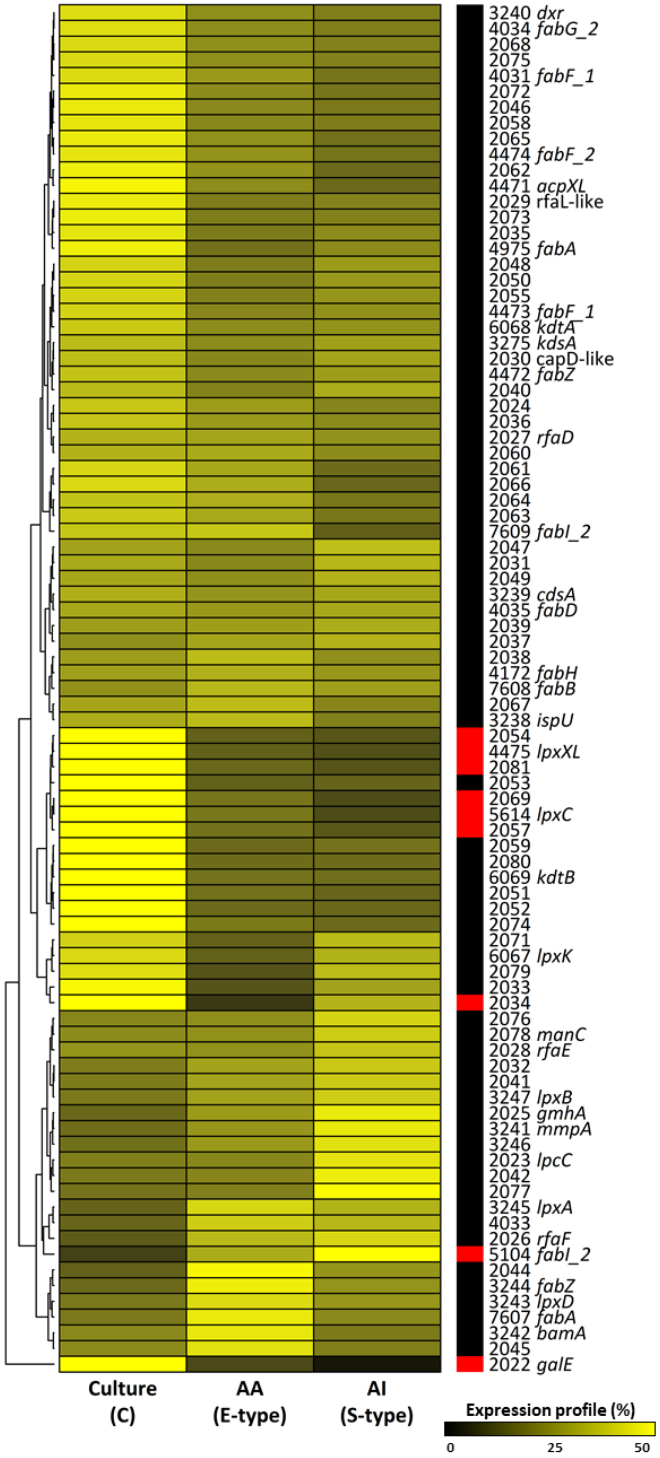


Fig S12. Expression of polysaccharide biosynthesis genes. Heatmap showing the expression of genes involved in polysaccharide (PS) synthesis. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

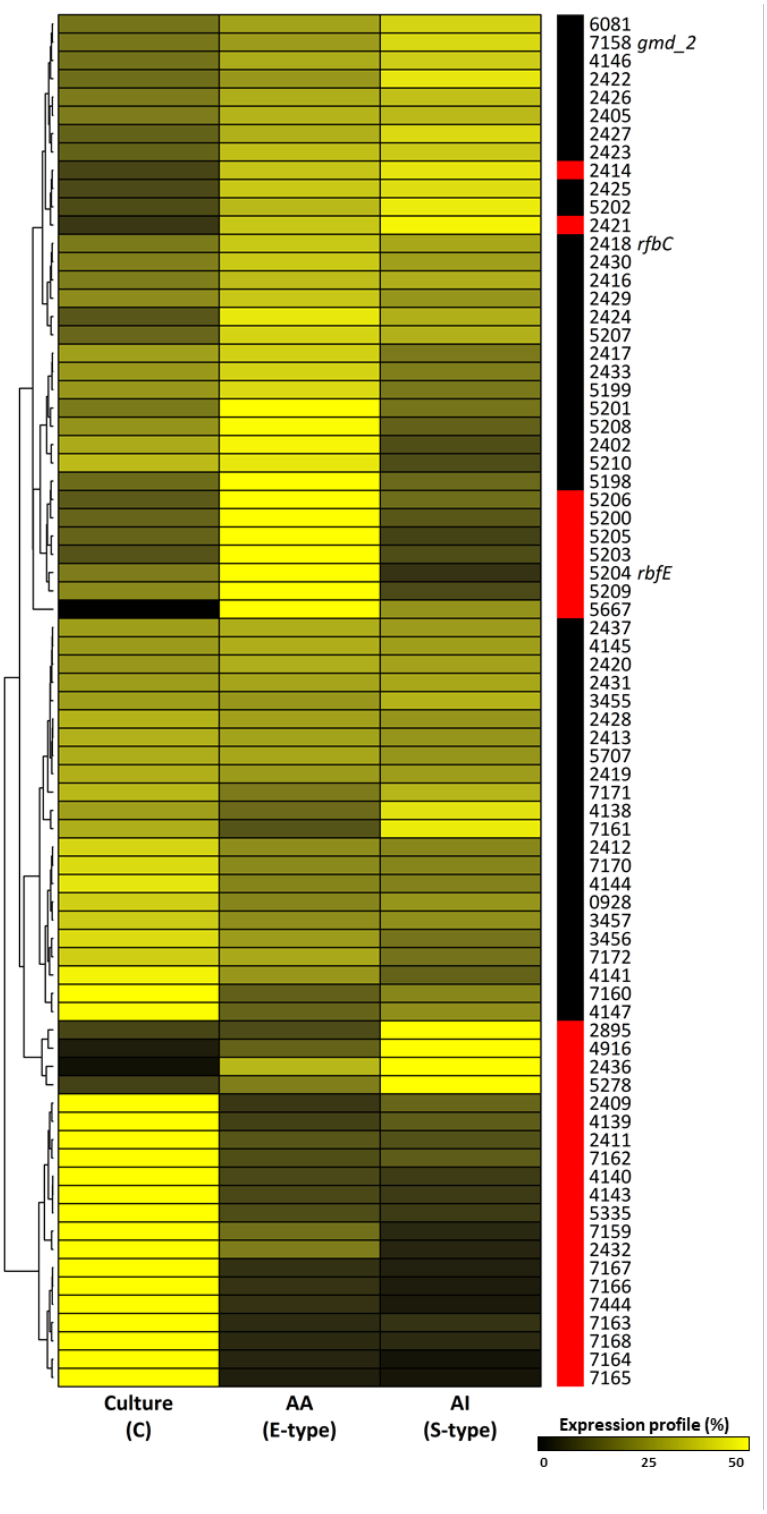


Fig S13. Expression of genes involved in chemotaxis, flagellum and pilus biosynthesis. Heatmap showing the gene expression of chemotaxis, flagellum and pilus biosynthesis genes. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

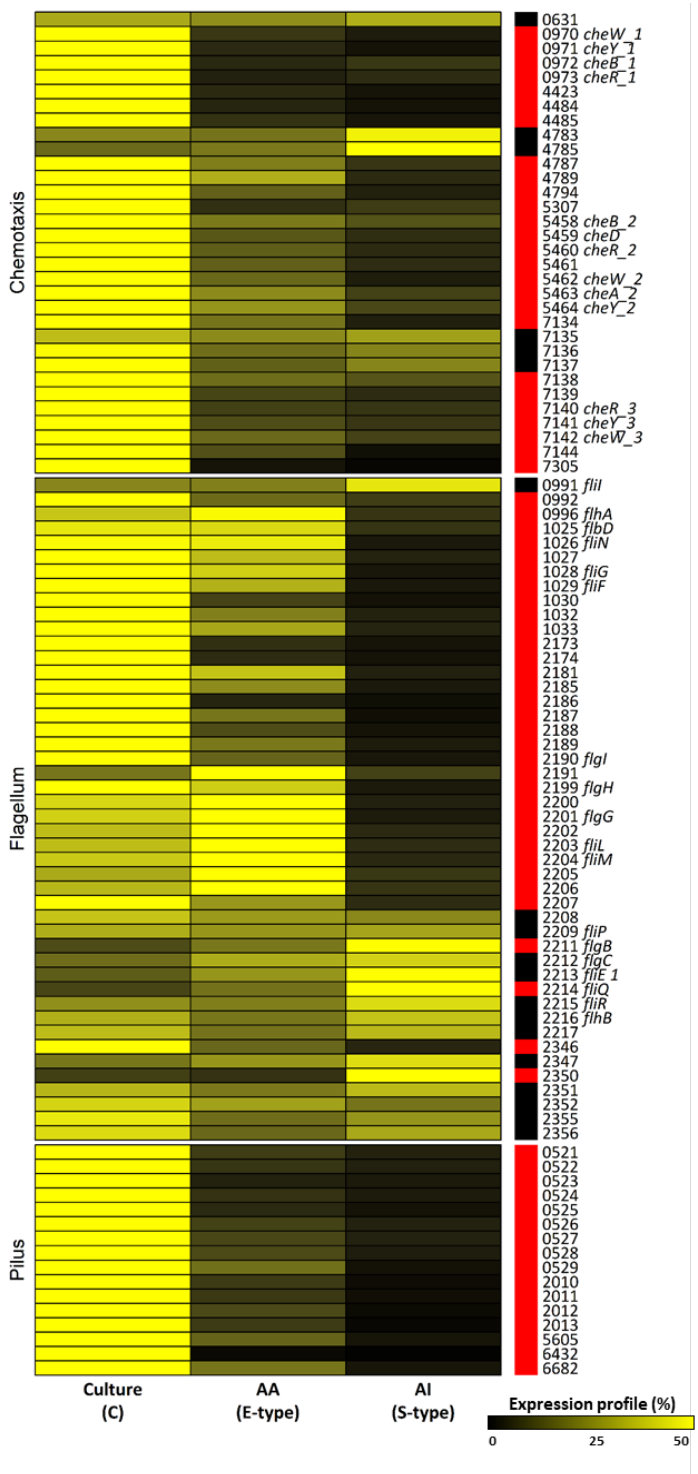


Fig S14. Expression of ribosomal genes. Heatmap showing the expression level of ribosomal protein encoding genes. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

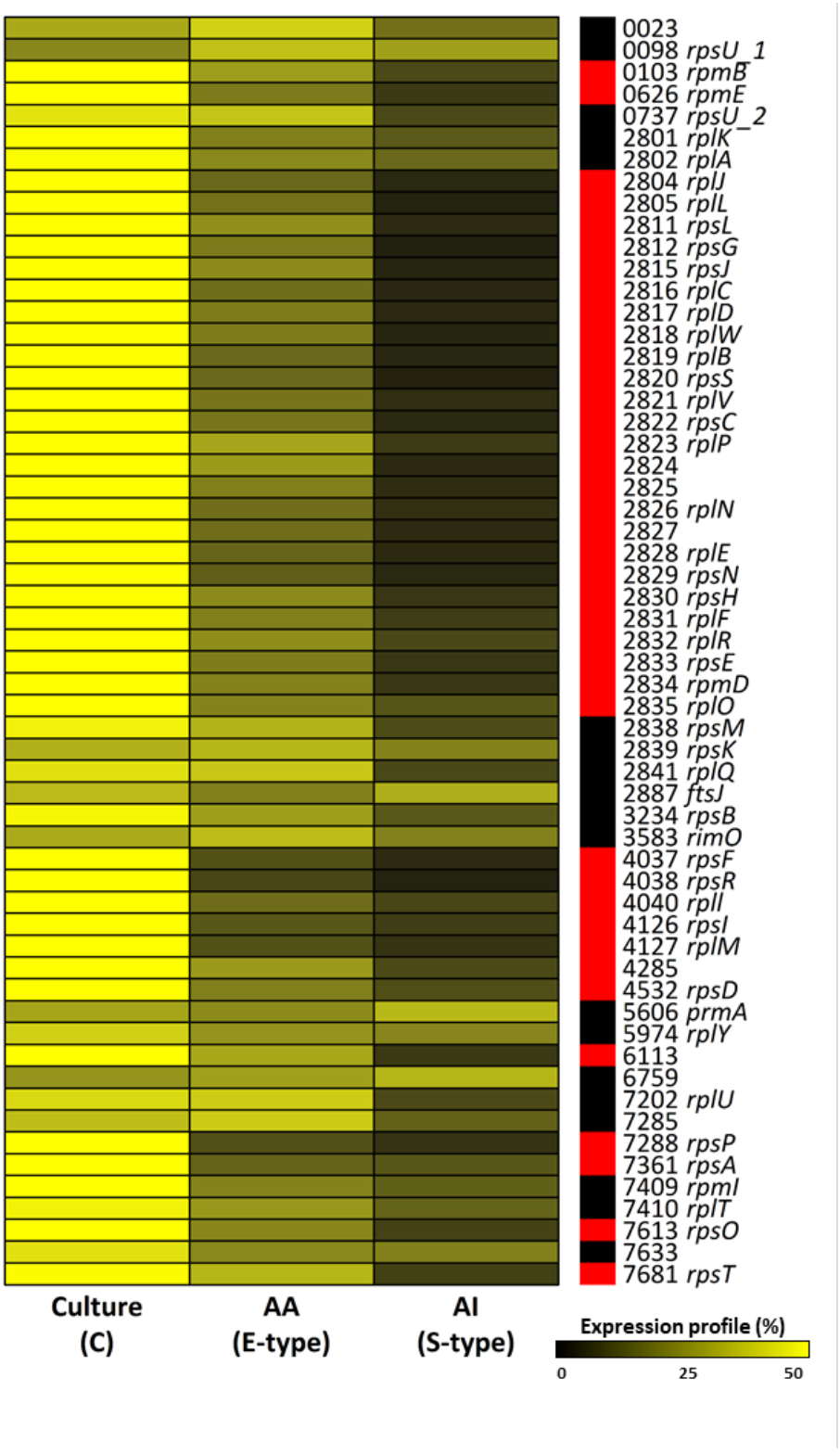


Fig S15. Expression of genes involved in bacterial photosynthesis. Heatmap showing the expression level of the genes involved in photosynthesis. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

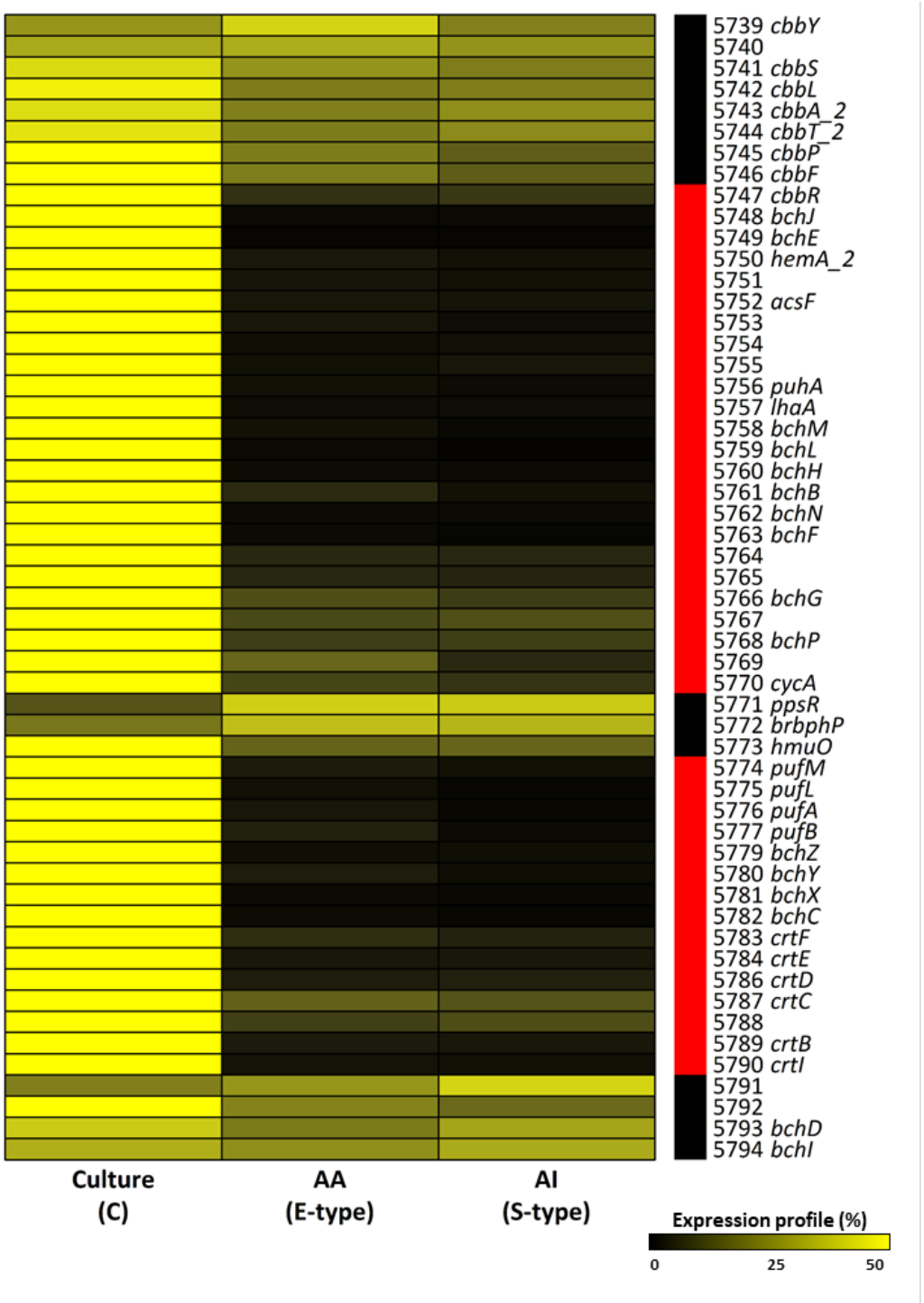


Fig S16. Expression of Non-Ribosomal Peptide Synthase (NRPS) and Quorum Sensing (QS) genes. Heatmap showing the expression of NRPS and QS encoding genes. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

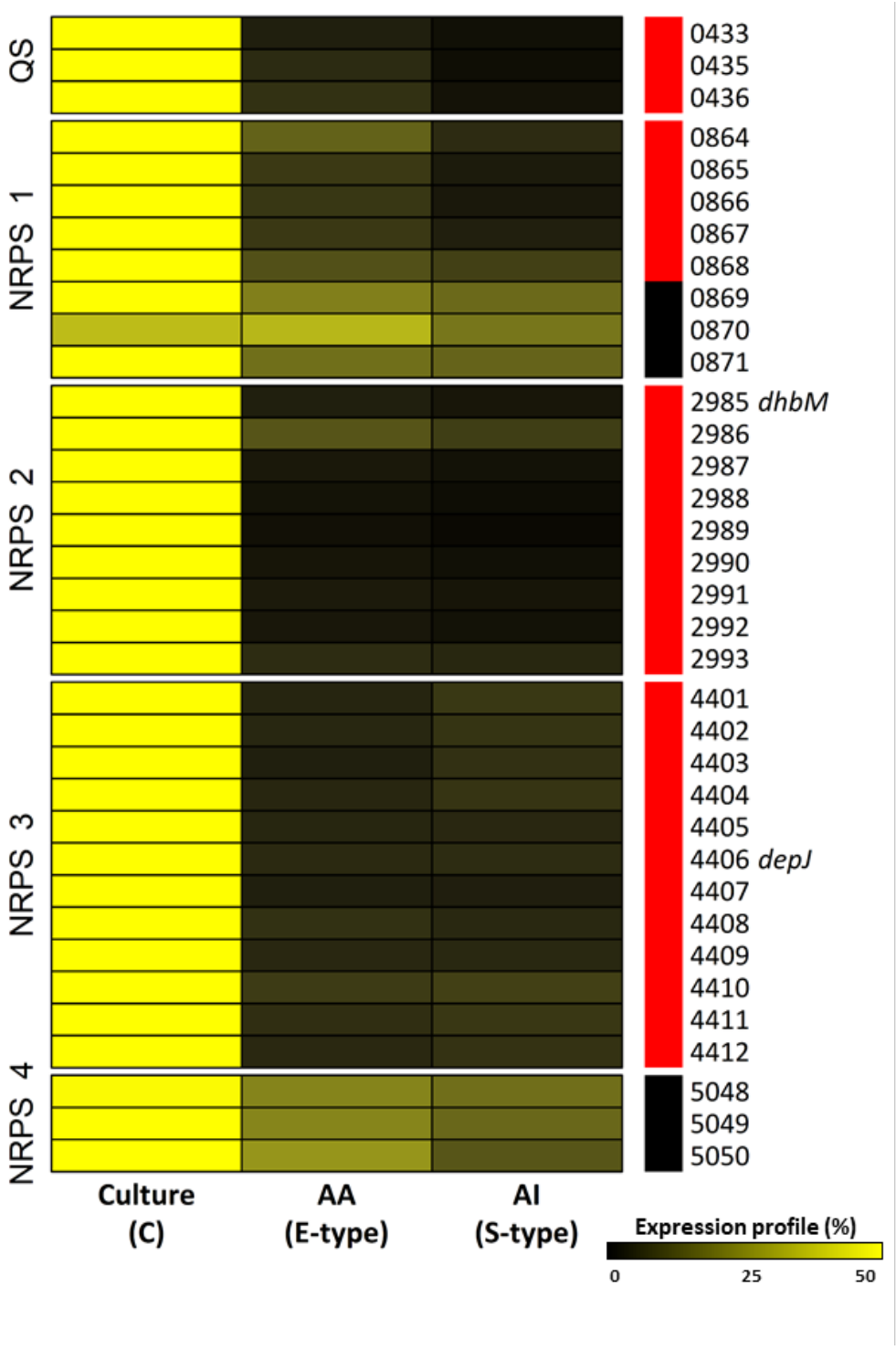


Figure S17. Expression of genes involved in biotin biosynthesis and utilization. Heatmap showing the expression of genes encoding biotin biosynthesis enzymes as well as biotin-dependent enzymes. The *bioADFB* genes are biosynthetic genes. The locus 5128-5135 carries genes encoding biotin dependent carboxylases. Other genes in the heatmap encode biotin-dependent carboxylases. The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

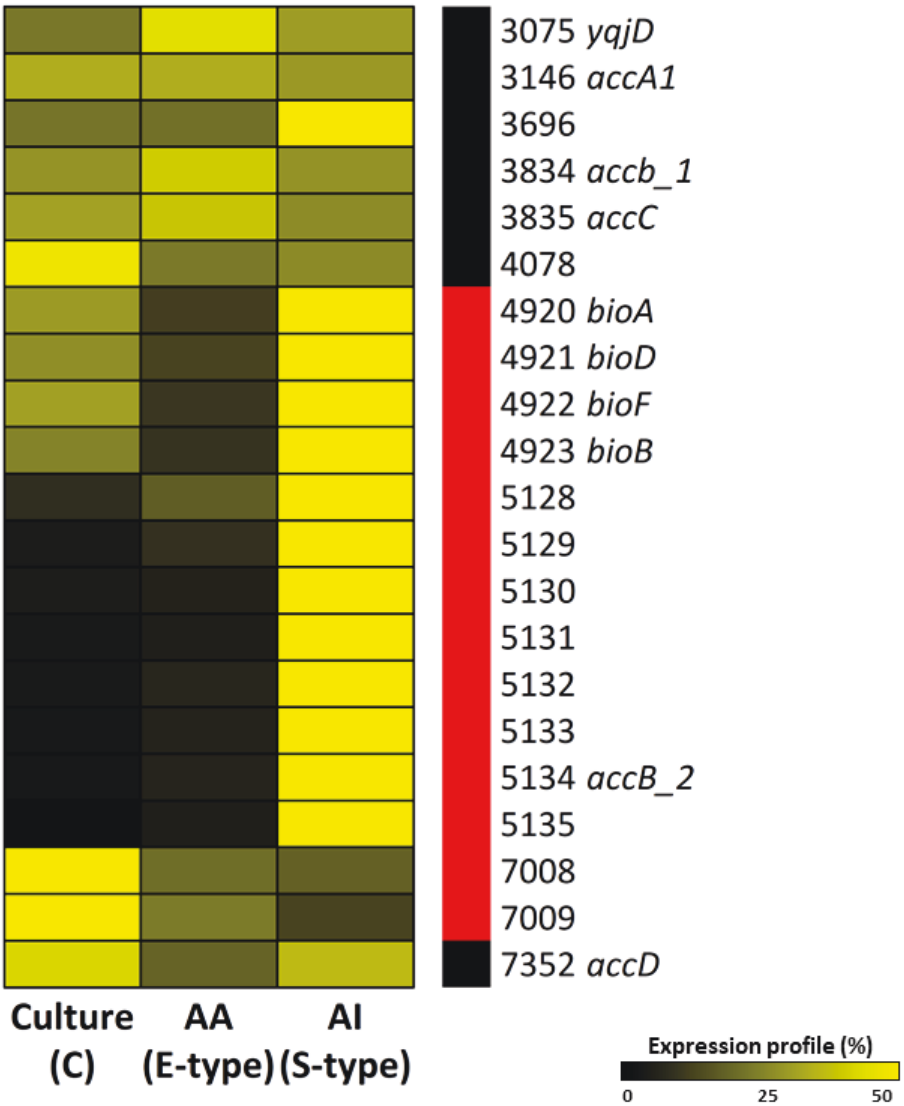


Figure S18. Expression of genes involved in vitamin B12 synthesis and utilization. Heatmap showing the expression of genes involved in B12 vitamin biosynthesis and of genes encoding the B12-dependent enzymes MetH (0497), RNR (3862) and MutAB (4651, 4653). The right column indicates whether the gene is a DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). The genes in rows are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2).

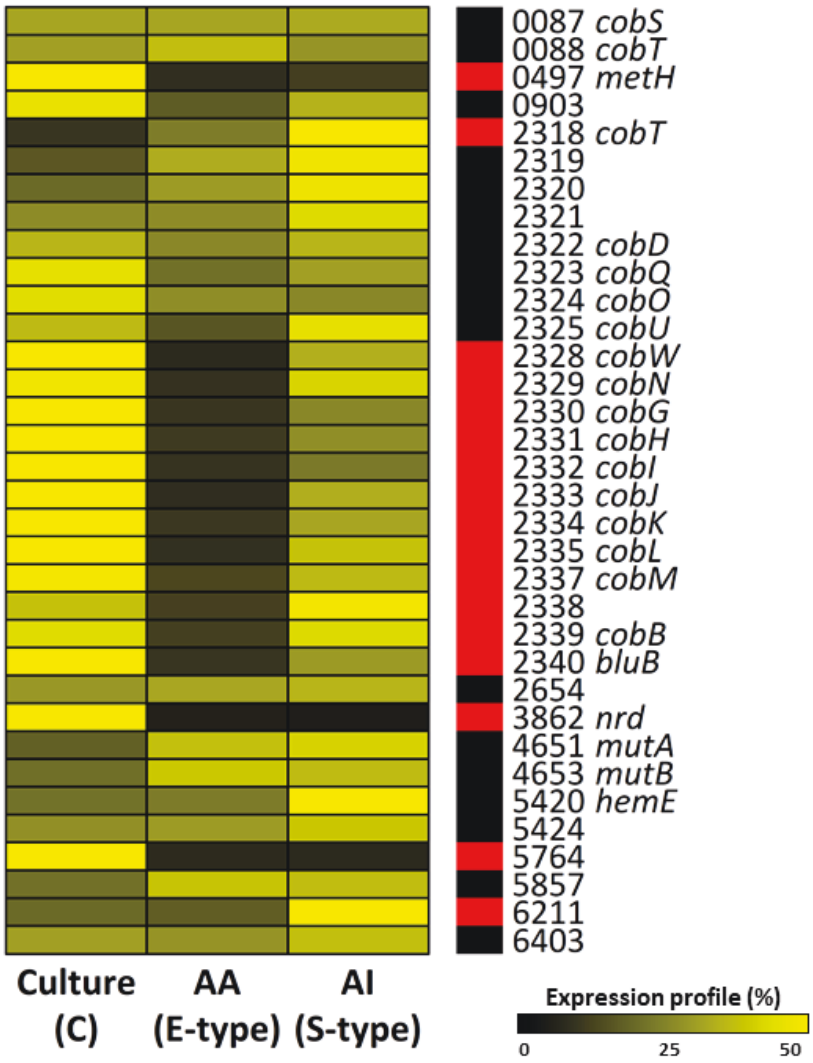


Figure S19. Nodulation phenotypes of 8 mutant strains. (A, B) Relative nitrogenase activity measured by the acetylene reduction assay of *A. afraspera* and *A. indica* nodules, respectively. (C, D) Nodule number per plant on *A. afraspera* and *A. evenia*, respectively. Letters are the results of Tukey post-hoc test ($p<0.05$). Red stars indicate significant differences compared to the WT strain. The genes inactivated in the mutants are identified by numbers that refer to the numerical value in the accession code BRAD285_v2_xxxx (Supporting Information Table S2). NI are non inoculated control plants.

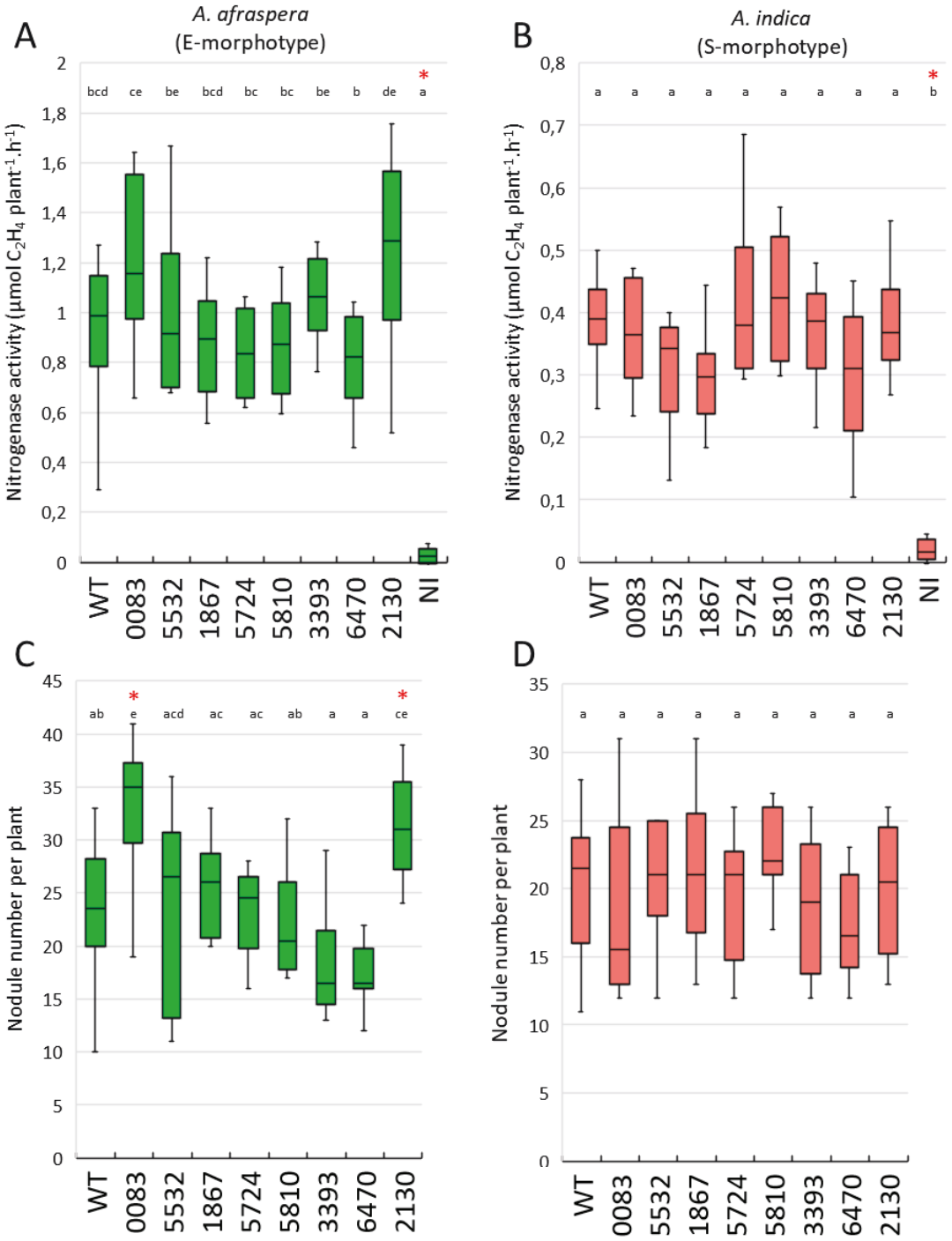
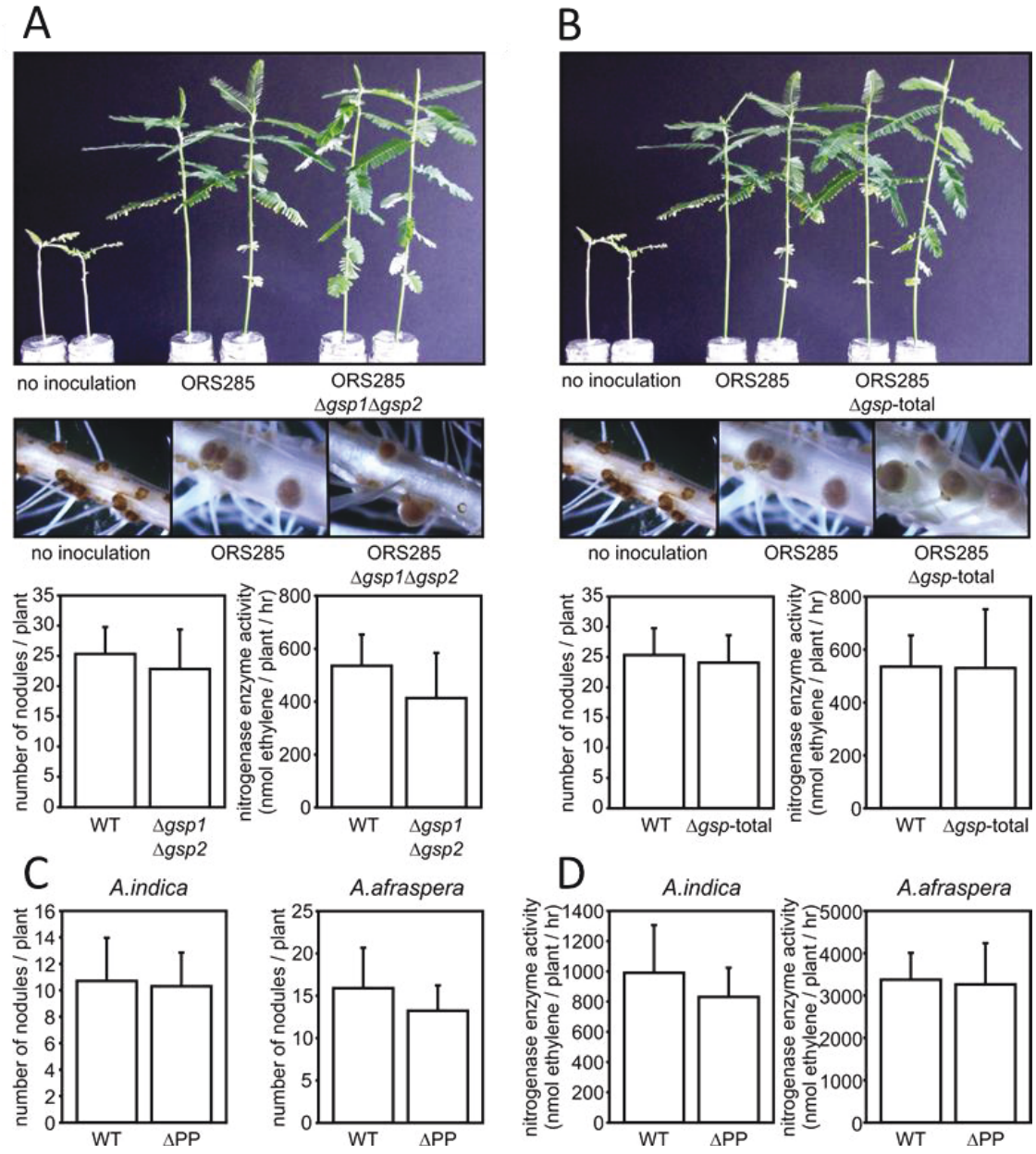


Figure S20. Nodulation phenotypes of mutants in the general secretion pathway. (A, B) Phenotypes of *Δgsp1Δgsp2* (double mutant of the individual *gsp* region deletions) and Δ*gsp*-total (deletion mutant with the complete locus removed) mutants in association with *A. indica*. Top panels show the growth of aerial parts of non-inoculated plants or plants inoculated with wild-type and mutant strains. The middle panels show root segments of non-inoculated plants or plants inoculated with wild-type and mutant strains. The lower panels show the quantification of nodule numbers and nitrogen fixation, measured by the Acetylene Reduction Assay. (C, D) Phenotypes of the prepilin peptidase deletion mutant (ΔPP, BRAD285_v2_1073) in association with *A. afraspera* and *A. indica.* Nodule number and nitrogen fixation rate have been measured in wild-type and ΔPP and display no significant differences.



3. From intracellular bacteria to differentiated bacteroids: transcriptome and metabolome analysis in *Aeschynomene* nodules using the *Bradyrhizobium* sp. ORS285 Δ*bclA* mutant

Florian Lamouche, Anaïs Chaumeret, Ibtissem Guefrachi, Olivier Pierre, Yves Dessaux, Florence Guérard, Françoise Gilard, Bertrand Gakière, Attila Kereszt, Peter Mergaert et Benoit Alunni

Avant-propos

Dans la continuité du chapitre précédent et toujours avant mon arrivée en M2, le jeu de données RNA-seq était en réalité également composé de la souche mutante Δ*bclA* (Guefrachi *et al.*, 2015). Les trois même conditions ont été générées, à savoir une culture libre et des bactéroïdes de *A. indica* et *A. afraspera* récoltés à 14dpi. Pour des raisons de simplification du discours, l'ensemble des jeux de données transcriptomiques a été scindé en deux études.

Le but initial de l'utilisation de ce jeu de données supplémentaire était double. Il visait d'abord à décrire le métabolisme et la physiologie des bactéroïdes incapables de se différencier dans les nodosités d'*Aeschynomene*. La possible fonction de ménage de la protéine BclA a aussi été investiguée en comparant les deux souches en culture libre.

Nous avons également utilisé la condition comme point de contrôle de l'établissement d'une interaction symbiotique, constituant un stade intermédiaire entre la vie dans le sol et la vie de fixateur d'azote intracellulaire. Ceci nous a alors permis de filtrer les gènes spécifiques de l'environnement végétal/intracellulaire mais non requis pour les processus de différenciation des bactéroïdes ou de fixation d'azote.

Nous avons combiné ici une approche de transcriptomique et de métabolomique sur nodosités complètes comparant les nodosités de *A. indica* et *A. afraspera* élicitées par la souche sauvage ou la souche mutante.

Comme précédemment, j'ai effectué l'ensemble des analyses de type RNA-seq, de l'étape de mapping à l'écriture du manuscrit. Un ancien post-doc (Olivier Pierre) a produit les échantillons de nodosités utilisés dans l'analyse métabolomique (collaboration Bertrand Gakière, IPS2).

Ce travail est en cours de révision dans la revue *Journal of Bacteriology*.

Summary

During legume-rhizobium symbiosis, free-living soil bacteria known as rhizobia trigger the formation of and infect plant root organs called nodules, where they adopt an intracellular lifestyle within nodule symbiotic cells and become nitrogen fixers known as bacteroids. Several legume lineages enforce their symbionts into an extreme cellular differentiation, comprising cell enlargement and genome endoreduplication, through the massive production of defensin-like peptides called Nodule-specific Cysteine Rich (NCR) peptides that are targeted to the bacteroids. In *Bradyrhizobium* sp. ORS285, which is differentiated in *Aeschynomene* spp. plants, a major bacterial determinant of this differentiation process is an antimicrobial peptide transporter called BclA, without which bacteroid cannot differentiate and fix nitrogen, thereby undergoing a rapid cell death within nodule cells. However, this absence of bacteroid differentiation occurs after infection, nodule organogenesis and bacterial internalization. The mutation of BclA could then be considered as a possible checkpoint in the establishment of the symbiotic interaction and constitutes an intermediate step between the life in the soil and intracellular life as nitrogen-fixing bacteroid. Here, we combined whole nodule metabolomics and bacterial transcriptomics to understand what the physiological consequences of the bacterial mutation on both partners are. Moreover, we discriminate sets of genes that require a functional BclA to be expressed and to drive a functional symbiosis, and others that rather seem to respond to the host microenvironment. Another of our major findings is that the Δ*bclA* mutant is constitutively expressing yet uncharacterized NRPS, suggesting possible functions of BclA outside the symbiotic process.

Introduction

Legume plants can fulfill their nitrogen requirements for growth thanks to a symbiotic interaction with soil bacteria called rhizobia. The bacteria are housed in dedicated plant organs called nodules (Udvardi and Poole, 2013). Rhizobia are released within host plant symbiotic cells by an endocytosis-like process and become nitrogen-fixing bacteroids during nodule formation. In several legume lineages, a terminal bacteroid differentiation (TBD) process occurs before effective nitrogen fixation, producing swollen polyploid bacteroids with altered viability (Mergaert *et al.,* 2006; Czernic *et al.,* 2015; Alunni and Gourion, 2016). The host plant controls TBD by targeting defensin-like antimicrobial peptides called Nodule-specific Cysteine

Rich (NCR) peptides to the bacteroids (Mergaert *et al.,* 2003; Van de Velde *et al.,* 2010; Czernic *et al.,* 2015). This process requires the DNF1 protein which is part of the signal peptidase complex that processes NCR peptides to allow their trafficking within vesicules up to the bacteroids (Wang *et al.,* 2010). This peculiar type of symbiotic antimicrobial peptides has been found in two distinct legume families, the Inverted-Repeat Lacking Clade (IRLC) (e.g *Medicago* genus) and the Dalbergioids (e.g *Aeschynomene* genus) (Montiel *et al.,* 2017; Czernic *et al.,* 2015). The NCR peptides of those two families are however not related and might have arisen by a convergent evolution process (Czernic *et al.,* 2015).

The successful completion of the TBD process also depends on bacterial factors, such as the BacA protein of *Sinorhizobium meliloti* (Glazebrook *et al.,* 1993). BacA is a transmembrane protein displaying similarities to the SbmA antimictobial peptide transporter of *Escherichia coli*. These proteins bare a SbmA_BacA domain (pfam05992), which corresponds to a subfamily of transmembrane domains found in ABC transporters, although SbmA and BacA lack the ATPase domain usually present in this class of ATP-dependant transporters. The Δ*bacA* mutant of *S. meliloti* does not differentiate and undergoes rapid cell death after its release inside wild type symbiotic plant cells (Haag *et al.,* 2011). However, when the plant mutant *DNF1* is inoculated by Δ*bacA* mutant, bacteroids fail to differentiate but remain alive (Wang *et al.,* 2010; Haag *et al.,* 2011). Thus, BacA appears to be required to limit the antimicrobial activity of the NCRs. Furthermore, homologs of BacA have been found in animal pathogens such as *Brucella abortus* and *Mycobacterium tuberculosis,* where BacA is required to establish a chronic infection of the host. Moreover, the fact that these BacA proteins can partially complement the Δ*bacA* mutant phenotype in *S. meliloti* suggest that BacA proteins are involved in facing a broader range of antimicrobial peptides (LeVier *et al.,* 2000; Domenech *et al.,* 2009; Wehmeier *et al.,* 2010; Arnold *et al.,* 2013).

Recent studies unveiled the importance of *Bradyrhizobium* sp. ORS278 and ORS285 BclA protein, a BacA-like transporter with an ATPase domain required for the establishment of a functional symbiosis with *A. afraspera* and *A. indica* (Bonaldi *et al.,* 2010; Guefrachi *et al.,* 2015). Analysis of the DNA content of wild-type (WT) and Δ*bclA* mutant bacteroids in both *Aeschynomene* plant hosts revealed that Δ*bclA* bacteroids fail to differentiate. These striking similarities between IRLC and Dalbergioids, in terms of recruited mechanisms both on the bacterial (BacA-like transporters) and plant sides (NCR-like peptides), suggest a convergent co-evolution process occurring at the plant-bacteria association level to tailor TBD in these distant legume clades and distant rhizobial species. Although nitrogen fixation is impaired in

Δ*bclA* mutant bacteroids, the intensity of the phenotype differs between the two bacteroid types. Indeed *A. afraspera* bacteroids are still fixing marginal amounts of nitrogen and bacteroid are all alive whereas *A. indica* bacteroids are completely unable to fix nitrogen and display higher membrane porosity and progressive cell death with mixed alive and dead bacteroid populations at 14 dpi (Guefrachi *et al.*, 2015). These observations suggest that although differences may exist in the peptide spectrum of the two host plants, the Δ*bclA* mutant can be seen in both hosts as an intermediate step between free living bacteria and functional bacteroids. Thus, the Δ*bcla* mutant can be used as a tool to dissect the process of bacteroid formation and to study the transcriptional response of bacteria in the plant environment before and after TBD and nitrogen fixation have taken place. In this study, we first analyzed the metabolome of *A. indica* and *A. afraspera* whole nodules elicited by *Bradyrhizobium* sp. ORS285 either WT (functional) or Δ*bclA* mutant (non-functional) strains. We then compared the transcriptomic profiles of *Bradyrhizobium* sp. ORS285 WT and Δ*bclA* mutant in exponential phase of culture and in mature nodules of *A. afraspera* and *A. indica*. Together, these data show to which extent the symbiosis is compromised and what is the effect of this phenomenon on the functioning of bacterial cells at the molecular level. Moreover, we use the Δ*bcla* mutant to discriminate bacterial functions that are expressed in response to the plant microenvironment once bacteria have entered the nodule versus the ones that are more presumably linked to the nitrogen fixing activity of differentiated bacteroids.

Material and Methods

Plant growth and inoculation

Bradyrhizobium sp. strain ORS285 wild type and Δ*bclA* mutant strain were grown at 30°C in YM medium supplemented with appropriate antibiotics (Guefrachi *et al.,* 2015). *A. indica* and *A. afraspera* plants were grown in transparent test tubes filled with BNM medium and inoculated with ORS285 wild type and mutant strains as described before (Lamouche *et al.,* 2018). Plants were grown at 28°C, 80% humidity and under a 16h : 8 h of light : dark regimen.

Metabolomic and elemental analyses

Fourteen dpi nodules were collected in liquid nitrogen and lyophilized. Metabolites and cofactors were extracted and analyzed by GC-MS and LC-MS respectively according to Su *et al.* (2016).

RNA extraction, transcriptomics and qRT-PCR

Nodules were harvested at 14 dpi, frozen in liquid N_2 and kept at -80°C until RNA extraction. For transcriptomic analysis, total nodule RNA was extracted as previously described (Mergaert *et al.*, 2003), and submitted to plant and bacterial rRNA and eukaryotic mRNA depletion using RiboZero kits for plants and for Gram-negative bacteria (Epicentre, Madison, USA). Total RNA of free-living cells was extracted from bacterial pellets obtained from cultures grown in YM medium at OD_{600}=0.5 using the hot phenol and hot SDS method (Chapelle *et al.*, 2015). Oriented libraries of bacterial mRNA were produced using the SOLiD total RNA-seq kit and sequenced on a SOLiD 5500XL (50bp per read). Each condition was performed in biological triplicates, resulting in the sequencing of 18 libraries which yielded ca. 30 million reads per sample, out of which ca. 9 million per sample were unique bacterial gene reads. The sequencing data of this project was submitted to the Sequence Read Archive (SRA) and obtained accession n° GSE108744 (SRA number: SRP128034) for the wild type strain and to be defined for the Δ*bclA* mutant strain.

For qRT-PCR analyses, total RNA was extracted using the Plant RNeasy kit (Qiagen). DNase treatment was performed on 2 µg RNA using the TurboDNA-free kit (Ambion, Life Technologies). Reverse transcription was performed with a RevertAid Reverse Transcriptase kit (Fermentas, Thermo Scientific). qRT-PCRs were carried out on a Roche LightCycler® 96 System using the Light Cycler Fast Start DNA Master SYBR Green I kit (Roche, http://www.roche.com/) and the primers listed in Lamouche *et al.*, (2018) for qRT-PCR experiments. All kits were used according to the manufacturer protocols. Cycling conditions were as follows: 95°C for 10 min, 50 cycles at 95°C for 5 sec, 58°C for 5 sec, and 72°C for 15 sec. The highly expressed 16S rDNA gene (rrs) and the moderately expressed BRAD285_v2_7630 gene annotated as miaB have been taken as references for qRT-PCR analyses. They have been chosen due to their invariant expression among conditions according to our transcriptome dataset, and their contrasted expression levels.

Statistical and differential expression analyses

CLC Workbench 10 software (CLC bio, Aarhus, Denmark) was used for trimming, sizing, and mapping the reads on the reference genome of Bradyrhizobium sp. ORS285 strain (Gully *et al.*, 2017). Only unique gene reads mapping on 80% of their length and displaying a 90% identity have been conserved for differential analyses. Count tables have then been filtered to retain only genes with a gene count over 1 count per million (cpm) in half of the samples of the dataset. Normalization and differential analyses were performed using generalized linear models as described in the DESeq2 package (version 1.12.4, Love *et al.*, 2014). Comparisons of gene expression values were performed using Wald-tests from DESeq2 package. The cutoff chosen for differentially expressed genes (DEG) are a False Discovery Rate (FDR) < 0.01 and a fold-change > 3 (log2 fold change (LFC) > 1.58). Principal component analysis of normalized data was computed using the FactoMineR package (v 1.33, Lê *et al.*, 2008). Heatmaps were produced using the ComplexHeatMap package (v 1.12, Gu *et al.*, 2016).The ratio comparing read repartition between the two strains in the three conditions of the experimental set-up, based on Cluster of Orthologous Genes (COG) database, has been calculated as follow: R i=log2[(∑ reads Δ*bclA* COGi)/ (∑ reads WT COGi * sf)], with sf being the size factor depending on RNA-seq library size.

Genome mining analyses

An antiSMASH query was performed on the *Bradyrhizobium* sp. ORS285 genome to identify possible enzymes involved in secondary metabolite synthesis, using the MiscroScope platform (Weber *et al.,* 2015; Vallenet *et al.,* 2017).

Results and discussion

Whole nodule metabolomics show differential nitrogen metabolism between hosts and reduced energetic metabolism in mutant nodules

We considered here Δ*bclA* conditions as controls of ineffective nitrogen fixation to determine nodule normal functioning by contrast with wild type strains (Guefrachi *et al.,* 2015). We followed the levels of 129 metabolites and nucleotide cofactors by respectively running

GC-MS and LC-MS analysis on whole nodules and obtained 47 significantly differentially (fdr<0.05 and significant post-hoc Tukey tests p<0.05) accumulated metabolites in 4 distinct classes as follows. The two first classes only consider the two WT conditions and comprise 27 and 7 metabolites differentially accumulated in *A. indica* and in *A. afraspera*, respectively. Comparing WT and Δ*bclA* nodules, 8 metabolites are accumulated in both WT conditions and 6 metabolites are specific for Δ*bclA* ineffective nodule conditions (Figure S1, Table S1).

In legumes, two main forms of fixed nitrogen can be exported from nodules to the aerial part of the plant. These can be metabolites such as asparagine (Asn), glutamine (Gln) and/or ureides (Oldroyd *et al.,* 2011). It seems that *A. indica* and *A. afraspera* have distinct strategies to transport fixed nitrogen to the other plant organs, with *A. afraspera* exporting both Asn and Gln whereas *A. indica* preferentially use Asn as exported form of nitrogen, as deduced from the accumulation of the corresponding metabolites in the symbiotic organs (Figure 1A).

Among the differentially accumulated metabolites in Δ*bclA* nodules, phosphoric acid and γ-aminobutyric acid (GABA) display higher concentrations in Δ*bclA* conditions than in wild type nodules (Figure 1B). Phosphate is often seen as a limiting factor of nitrogen fixation and the accumulation of phosphoric acid could be due to an absence of its consumption by non-functional bacteroids (Sulieman *et al.*, 2013). Indeed, differentiated bacteroids of ORS285 strain show a high expression of genes involved in phosphate and phosphonate import in wild type bacteroids (Lamouche *et al.*, 2018). Then, Δ*bclA* bacteroids do not seem to be starved in phosphate by the host plant. On the other hand, GABA accumulation has been observed as potentially inducing nitrogen fixation in *Medicago truncatula* in case of nitrogen stress and plants treated by GABA show induced nitrogen fixing activity (Sulieman and Schulze, 2010; Sulieman, 2011). Then, *Aeschynomene* spp. could either accumulate GABA in nodules elicited by Δ*bclA* strain to boost deficient nitrogen fixation, or GABA accumulation would reflect the nutritional stress induced by the strong reduction/absence of nitrogen fixation in mutant nodules (Sulieman, 2011). However, as proposed in the *Pisum sativum-Rhizobium leguminosarum* association, GABA could also be an amino-acid source for bacteroids permitting a decarboxylation shunt in the Krebs cycle (Prell *et al.*, 2009). This shunt could be advantageous for the plant, avoiding the possibility for bacteria to induce polyhydroxybutyrate (PHB) synthesis as carbon storage, thereby avoiding a carbon loss for the plant and ensuring that the flow of carbon is used for nitrogen fixation in the bacterium. Differentiated bacteroid could then be fed by the plant with GABA, and its accumulation in nodules elicited by Δ*bclA* strain could be due to its non-consumption by non-functional bacteroids.

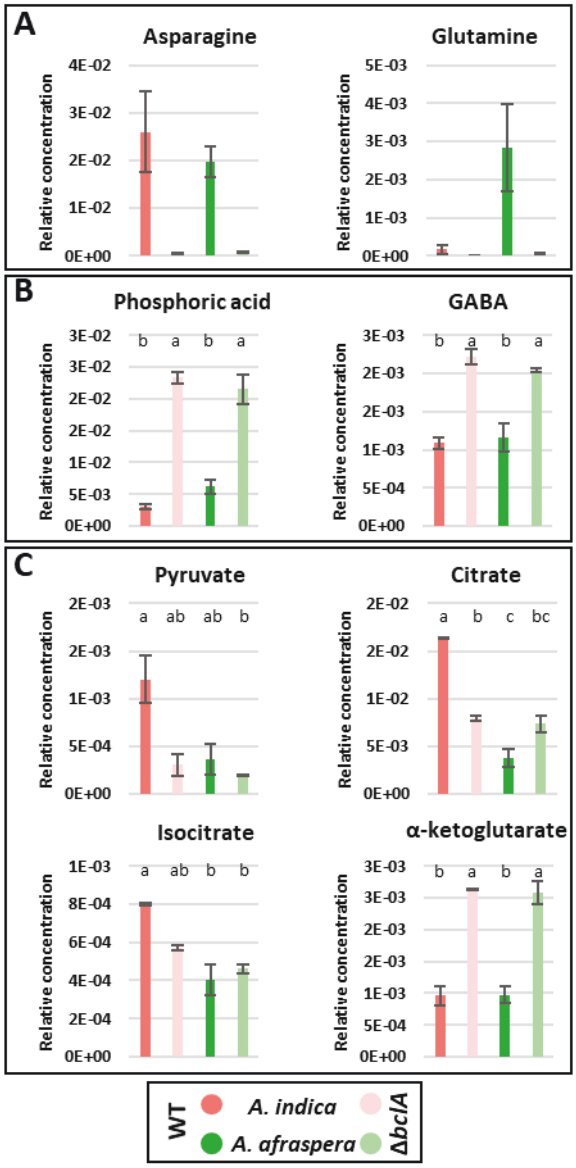


Figure 1. Whole nodule metabolomics show significant differential accumulation of several metabolites in nodules elicited by Δ*bclA* bacteria. (A) *A. indica* and *A. afraspera* export distinct forms of nitrogen from nodules. (B) Nodules elicited by *ΔbclA* strain accumulate phosphoric acid and γ-aminobutyric acid (GABA) respectively suggesting limited phosphate uptake in bacteroids and plant nitrogen stress. (C) *A. indica* nodules display a higher primary metabolic rate leading to glutamate formation suggesting a higher nitrogen fixation rate. Data are mean (n = 2) ± SE (Standard error). Statistical significance was tested with One-Way ANOVA (fdr ≤ 0.05) and post-hoc Tukey tests where letters indicate significant differences.

Focusing on the Krebs cycle, *A. indica* nodules seems to have a higher level of metabolites upstream from glutamate conversion compared to *A. afraspera*, such as pyruvate, citrate and isocitrate (Figure 1C). Glutamate is produced by the amination of α-ketoglutarate in the plant symbiotic cells using ammonia produced by the bacteroids (Figure 1C; Oldroyd *et al.,* 2011). Consequently, *A. indica* nodules seem to have a higher primary metabolic rate probably linked to glutamate production and thus a higher nitrogen fixation rate, although fluxomic analysis would be required to confirm these suppositions deduced from the static view provided by whole nodule metabolomics. Δ*bclA* elicited nodules on the contrary display high accumulation of α-ketoglutarate, probably due to the absence of nitrogen fixation (Figure S2). Furthermore, while wild-type bacteroids display high metabolic activities as deduced form the pools of

ATP/ADP/AMP/NAD/NADP measured by LC-MS, we observed slightly higher levels of those metabolites in *A. indica* bacteroids (Figure. S1). These metabolic features support the previous observations of the higher symbiotic efficiency of S-type bacteroids compared to *A. afraspera* E-type bacteroids and provide new insights on the possible underlying mechanisms (Lamouche *et al.,* 2018). Deletion of *bclA* triggers a drastic reduction of those same metabolites suggesting a much lower metabolic rate of the whole nodules.

After this initial description of nodule metabolic dysfunction in the mutant nodules, we then wanted to investigate the physiology of non-differentiated Δ*bclA* bacteroids by performing a transcriptomic study.

General description of the transcriptomic dataset

We recently published the analysis of the WT transcriptome of *Bradyrhizobium* sp. ORS285 in symbiosis with *A. afraspera* and *A. indica* (Lamouche *et al.,* 2018). Additionally, samples corresponding to the Δ*bclA* mutant were produced simultaneously and the comparative analysis of these genotypes is presented here. The transcriptomic dataset (Table S2, GEO: to be defined) is thus composed of the two bacterial genotypes (ie. WT and Δ*bclA*), each in two bacteroid (ie. A. afraspera and A. indica at 14dpi) and one culture conditions, all produced and analyzed as biological triplicates (Figure 2A). Principal component analysis (PCA) revealed that Δ*bclA* bacteroids have a very different transcriptome compared to the WT bacteroids. Indeed, a clear partitioning of the samples was observed, with the first axis of the PCA (42% of the observed variance) separating the culture from the bacteroid conditions, the second axis (17% of the observed variance) separating the Δ*bclA* from WT bacteroids, and the third one (12% of the observed variance) separating A. indica and A. afraspera hosts (Figure 2B&C). The validity of the transcriptomic dataset has been assessed by qRT-PCR on a subset of 23 genes differentially regulated (Lamouche *et al.*, 2018). The log2-fold changes (LFC) obtained by both techniques were compared and appeared to be consistent (linear regression R2 value = 0.75; Figure S3). A general comparison of the transcriptomes was obtained by a ratio of read repartition between WT and Δ*bclA* conditions using Cluster of Orthologous Genes (COG) classification (Figure. S4). This COG analysis for the WT strain only revealed that bacteroids display an upregulation of genes involved in energy production/conversion (COG class C) and inorganic ion transport/metabolism (COG class P) on the one hand and by the drastic downregulation of genes

involved in translation (COG class J), signal transduction (COG class T), secondary metabolites biosynthesis (COG class Q) and cell motility (COG class N) on the other hand (Lamouche *et al.*, 2018). Contrastingly, Δ*bclA* bacteroids display an opposite transcriptional response compared to the WT, with a downregulation of class C and P, and an upregulation of all the other COG classes, showing the absence of the nitrogen-fixing metabolic specialization undergone by the mutant bacteroids (Figure. S4). The most interesting COG classes observed are defense mechanisms (COG class V) which are upregulated in Δ*bclA* bacteroids and the secondary metabolites biosynthesis (COG class Q) which is upregulated in the three Δ*bclA* conditions. More precisions concerning whole-genome differentially expressed gene regions are brought by a circular representation of the three comparisons of interest on the physical map of the bacterial chromosome (Figure 2D). The impact of the Δ*bclA* mutation on the bacterial transcriptome is highest in A. indica, with 962 differentially expressed genes (DEG, see Methods for chosen cutoffs) between the two bacterial genotypes, followed by A. afraspera, symbiosis with 270 DEG and only 114 DEG between the two genotypes in free-living culture.

Deletion of the *bclA* gene has a very limited impact on bacterial gene expression in culture

Apart from later discussed non-ribosomal peptide synthase (NRPS) expression which is highest in mutant bacteria, *bclA* deletion has a very limited effect on the transcriptome of free-living bacteria. Notably, Δ*bclA* mutants have an increased expression of a putative LrgAB two component system (BRAD285_v2_0350/0351), involved in the regulation of peptidoglycan synthesis. Mutant bacteroids also display an induction of genes encoding two isoquinoline 1 oxidoreductases (iorA_1 and BRAD285_v2_4828/4829), isoquinoline being a precursor of alkaloid antimicrobial compounds. *bclA* deletion appears to downregulate hydrogenase encoding genes such as *hypCDEF*/*hoxA*. Similarly, genes encoding a cytochrome C oxidase *ccoOPQS* are downregulated.

Using the absence of BclA to better define nitrogen fixation genes

Our previous study on wild type *Bradyrhizobium* sp. ORS285 showed 1342 DEG between bacteroid and culture conditions and more specifically 679 genes upregulated in bacteroid conditions in both hosts. Eleven of them were selected for mutagenesis and the corresponding mutants did not display any symbiotic phenotype (Lamouche *et al.,* 2018).

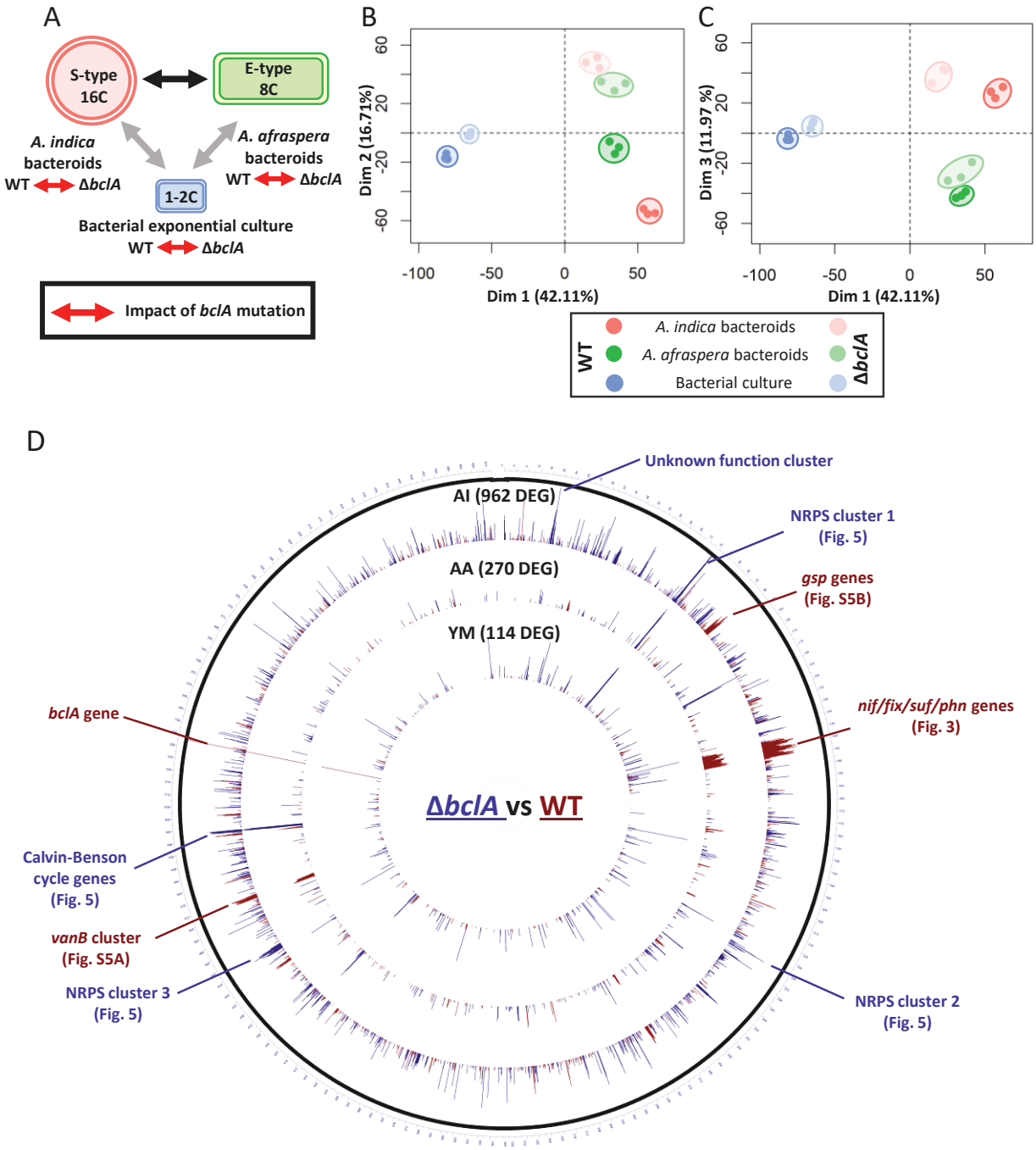


Figure 2. Experimental setup and general overview of the dataset. A. Experimental setup displaying the 6 biological conditions and the comparisons of interest. Bacteroid specific expression patterns (gray arrows) and host specific genes (black arrow) of the wild type bacterium have been analyzed in Lamouche *et al.*, 2018. (B, C). Principal component analysis displaying an excellent separation of the 6 conditions and repeatability of biological replicates. D. Circular representation of *Bradyrhizobium* sp. ORS285 chromosome displaying the three comparisons between WT and Δ*bclA* strains. Blue and red peaks display upregulated and downregulated genes in Δ*bclA* condition compared to the WT strain, respectively. AI: *Aeschynomene indica* bacteroids; AA: *Aeschynomene afraspera* bacteroids; YM: Free-living culture.

The phenotype of Δ*bclA* mutant resulting in aborted bacteroid differentiation and reduced/absence of nitrogen fixation makes it a suitable tool as a reference condition of intracellular bacteria for transcriptomics. It allows to filter genes responding to the plant nodule microenvironment/endosymbiotic lifestyle but that are not required for the differentiation and nitrogen fixation processes. Here we focused our analysis on the 679 bacteroid-specific genes.

Thus, we separated these 679 genes in seven gene categories based on their differential expression between Δ*bclA* and WT bacteroid conditions (Figure 3A, Table S3). The first class includes a list of 10 genes upregulated in both Δ*bclA* bacteroid conditions as compared to the wild type situation (Figure 3B). These gene mostly encode proteins of unknown function and the most interesting gene cluster, BRAD285_v2_4913-4916, encodes proteins involved in glycogen metabolism and an ABC transporter. This cluster is of particular interest because its expression is both bacteroid specific and upregulated in Δ*bclA* bacteroids compared to the WT ones. Moreover, *S. meliloti* ortholog of BRAD285_v2_4916 is upregulated when bacterial cultures are treated by NCR peptides (Tiricz *et al.*, 2013). Thus, we could speculate that this gene is involved in NCR peptide response, which could be even higher when BclA is absent, thereby increasing the NCR effect on the bacterial envelope. More importantly, *S. meliloti* ortholog of this gene is called *ndvA*. It has already been characterized and encodes a β1-2 glucan transferase required for an effective infection process in *M. sativa* nodules, but also for crown-gall tumor formation by *Agrobacterium tumefaciens* (Dylan *et al.*, 1986, Stanfield *et al.*, 1988).

The second class comprises 97 genes that are drastically downregulated in Δ*bclA* bacteroids, among which is the genomic region containing *nif*, *fix*, *suf*, *mop* genes (but not *phn* genes) which is very highly induced in wild type bacteroid conditions, representing more than 20% of the transcripts in wild type bacteria (Lamouche *et al.,* 2018). This gene set is thus the most promising to define candidate genes with a potential role in nitrogen fixation and bacteroid differentiation processes. Those genes apparently require BclA protein as a checkpoint to be expressed (Figure 3B, Figure S5A). The *vanB* cluster which is specific to photosynthetic bradyrhizobia and highly induced in WT bacteroids is also repressed in mutant bacteroids. This suggests an important role of one or some genes of this cluster in *Bradyrhizobium-Aeschynomene* symbiosis even if a *vanB* insertion mutant does not display any symbiotic phenotype (Figure 3B, Figure S5A, Lamouche *et al*., 2018). Apart from those two genomic regions, this gene selection unveils three new gene clusters that could possibly be required for bacteroid differentiation and/or nitrogen fixation. The first one (BRAD285_v2_0890-0891) encode a DnaJ-like protein and a putative ferredoxin respectively. Similarly, the second one

(BRAD285_v2_3886-3889) also encodes a ferredoxin and proteins of unknown function. The third one (BRAD285_v2_5727-5730) displays a very poor gene annotation. These loci have not been mutated yet, but it could be interesting to check their requirement for the establishment of a functional symbiosis. On the other hand, the other gene clusters are less good candidates to explain the nitrogen fixation or differentiation failures in the Δ*bclA* mutant, as they do not require a functional BclA to be expressed during symbiosis. The expression of these genes could be associated with plant environment specificity and/or intracellular lifestyle (Figure 3B). This is the case of *phn* and *pstSCAB*/*phoUB* genes that do not display the same drastic downregulation. Similarly, the type II secretion systems (*gsp*) cluster is not repressed in Δ*bclA* bacteroids either (Figure 5SB). This suggests a possible role of this secretion system in plant colonization but not a requirement for nitrogen fixation and bacteroid differentiation in ORS285, then explaining the absence of symbiotic phenotype in the corresponding mutants (Lamouche *et al.,* 2018). In this gene category can also be found *phyR*-σ^{EcfG} regulatory genes involved in general stress response in α-proteobacteria. They are known to display a symbiotic phenotype in *Glycine max* nodules elicited by *Bradyrhizobium diazoefficiens* USDA110 *phyR* and σ^{EcfG} mutant strains (Gourion *et al.*, 2009, Lederman *et al.*, 2018). However, their expression pattern suggest that these genes do not have a specific role in nitrogen fixation, but are rather required for the early steps of colonization, infection and nodule development (Lederman *et al.*, 2018). Moreover, the symbiotic phenotype cannot be related to bacteroid differentiation insofar as *G. max* does not produce NCR peptides. Furthermore, *nifA* transcription factor is induced in mutant bacteroids compared to the WT, especially in *A. indica* nodules. This observation is a bit surprising because *nifA* expression is supposed to induce the expression of nitrogen fixation genes which are completely repressed in Δ*bclA* bacteroids (Figure S5A). This result could be explained by a high expression of *nifA* gene but NifA could be inactivated by O_2 in non-functional leghemoglobin-poor nodules such as the ones elicited by Δ*bclA* mutant (Guefrachi *et al.*, 2015 ; Barrière *et al.*, 2017). In *A. indica* nodules elicited by Δ*bclA* strain, *napABCDE* genes encoding a nitrate reductase (first step of denitrification) are upregulated and have been shown to be under the control of *nifA* in *B. diazoefficiens* USDA110 (Bueno *et al*., 2010). As a consequence of their inability to differentiate and fix nitrogen inside *A. indica* nodules, Δ*bclA* bacteroids might require the use of nitrate reductase activity explaining the upregulation of *nifA* followed by the one of *napABCDE* genes.

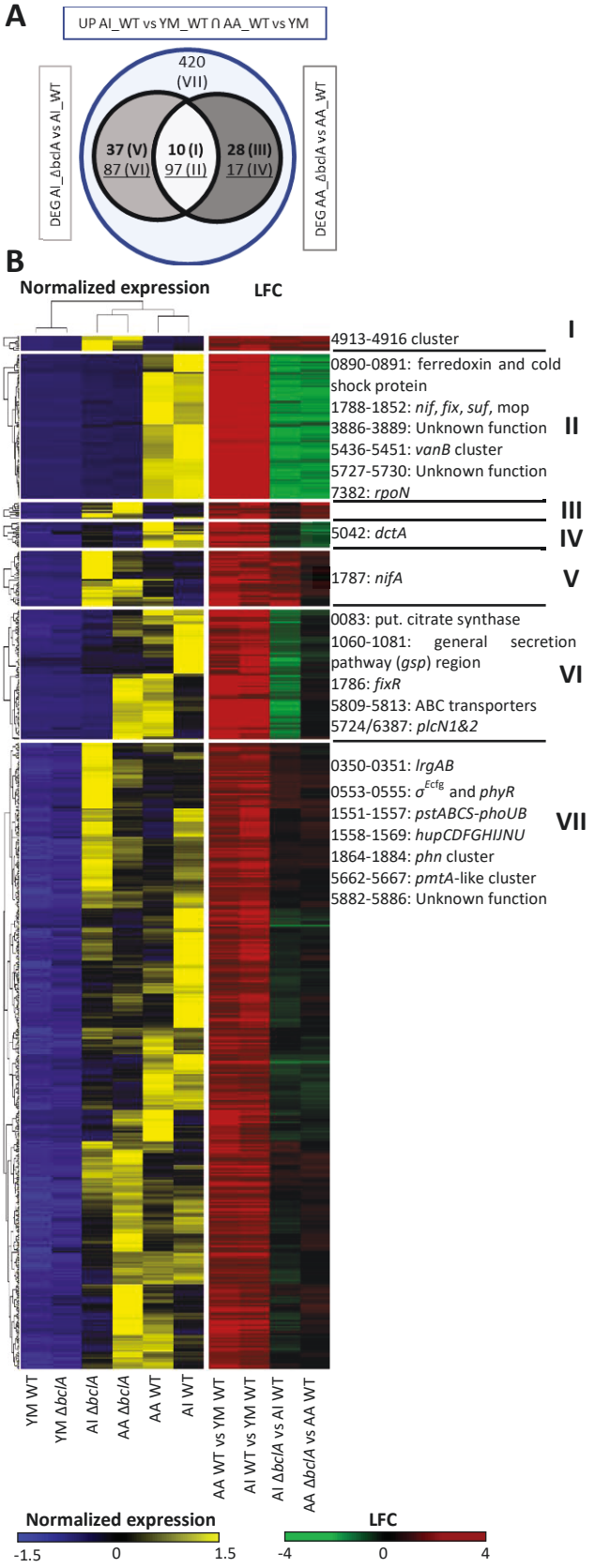


Figure 3. Repartition of the bacteroid specific genes in Δ*bclA* bacteroid conditions. A. Euler diagram showing DEG gene repartition of the 679 differentially expressed genes between Δ*bclA* and WT bacteroid conditions (fdr<0.01 and |LFC|>1.58) among the upregulated bacteroid geneset (upregulated in E- and S-type bacteroids as compared to the culture reference defined in Lamouche *et al.*, 2018). Bold and underlined numbers are respectively upregulated and downregulated in given conditions. B. Heatmap and hierarchical clustering separated in 6 classes defined by (A) of the differentially expressed genes in both Δ*bclA* and WT bacteroid comparisons. The color-coded scale bars for the normalized expression and LFC of the genes are indicated below the heatmap. Gene accessions are represented by the numerical value of their BRAD285_v2_xxxx format. AI: *A. indica*; AA: *A. afraspera*; YM: culture condition.

Using the absence of BclA to filter genes that are not essential to host adaptation for nitrogen fixation and bacteroid differentiation

Similarly, we previously defined 515 host-specific genes potentially involved in host adaptation and/or bacteroid higher differentiation and symbiotic efficiency observed in *A. indica* bacteroids (Czernic *et al.*, 2015 ; Lamouche et *al.*, 2018). Here again, the Δ*bclA* bacteroid conditions could be used to filter host specific functions that are simply triggered by the plant environment/endosymbiotic lifestyle but not essential to host adaptation for nitrogen fixation and bacteroid differentiation. Indeed, genes specifically induced in a host plant but downregulated in the same Δ*bclA* bacteroid condition are more likely to be great candidate genes for effective host-specific bacteroid adaptation. First, we selected host-specific genes that are upregulated in the same host condition compared to the reference culture to get rid of free-living specific genes. Then we used k-means clustering to distinguish 5 genes classes of interest (Figure 4, Table S4). The first class contains *A. indica* specific genes that are repressed in the same Δ*bclA* bacteroid condition. We found out some already described genes such as a putative citrate synthase (BRAD285_v2_0083, one of the most highly expressed gene in WT bacteroids), a cluster of ABC transporters (BRAD285_v2_6210-6215) whose insertion mutation unfortunately displayed no symbiotic phenotype (Lamouche *et al.*, 2018). Moreover, two siderophore receptor proteins (BRAD285_v2_1686-1687 and BRAD285_v2_5331-5332) are S-type specific and downregulated in Δ*bclA* bacteroids ; making them good candidate genes for functional analyses (Figure 4).

On the other hand, another class of *A. indica* specific genes contains functions that are not differentially expressed between the two bacterial genotypes in *A. indica*, such as *kdpABC* cluster and the poorly annotated BRAD285_v2_2893-2895 cluster probably involved in cellulose synthesis. The induction of these genes might rather be triggered by *A. indica* cell microenvironment rather than being linked to host adaptation for an effective symbiotic interaction, explaining the absence of symbiotic phenotype observed for *kdpA* and *kdpB* insertion mutants (Lamouche *et al.*, 2018). The *kdpABC* cluster however, is surprisingly upregulated in *A. afraspera* Δ*bclA* bacteroids. Similarly, the third and fourth classes are *A. afraspera* specific genes with very low variation between the two genotypes in *A. afraspera* nodules. These genes have probable very little effect on the symbiotic process either. The last gene class is the one of *A. afraspera* specific genes downregulated in Δ*bclA A. afraspera* bacteroids. This is the case of *asd* (aspartate-semialdehyde dehydrogenase), *aroG* (shikimate kinase) and we unfortunately already pointed them out in our previous study and tested for

functional studies with no symbiotic phenotype (Lamouche *et al.*, 2018). A poorly annotated gene cluster, BRAD285_v2_2620-2626 also belongs to this gene class, as well as the malate transporter encoding gene *dctA*, already described in various studies as one of the main entry routes for plant carbon in the symbiotic bacterium (White *et al.*, 2007). Apart from monitoring the behaviour of already described genes to better define the possible nature of their expression pattern, *bclA* mutation triggers in bacteroids the induction of other genes of interest.

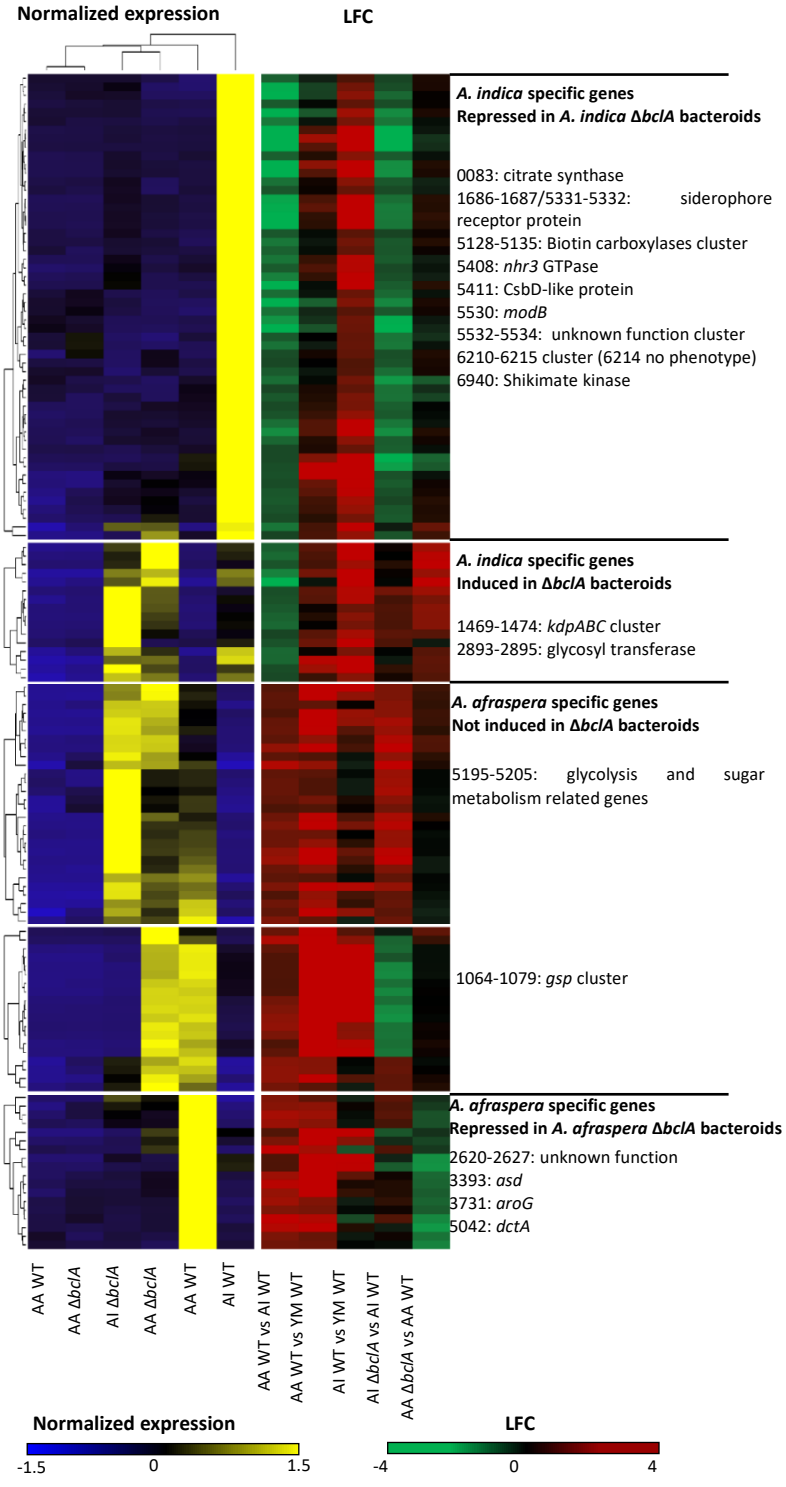


Figure 4. Heatmap of the 135 differentially expressed genes between E-type and S-type bacteroids at least differentially expressed in one bacteroid comparison between Δ*bclA* and WT conditions. Cutoffs are fdr<0.01 and |LFC|>1.58. The color-coded scale bars for the normalized expression and LFC of the genes are indicated below the heatmap. Gene accessions are represented by the numerical value of their BRAD285_v2_xxxx format.

bclA deletion triggers the expression of non-ribosomal peptide synthetases

Bradyrhizobium sp. ORS285 can produce several secondary metabolites. Analysis with the antiSMASH algorithm revealed the presence of one N-acyl homoserine lactone (AHL) synthesis cluster probably producing a quorum sensing signal and four non-ribosomal peptide synthetase (NRPS) gene clusters determining the production of so far uncharacterized products. NRPS are multimodular enzymes, with every module requiring three domains for peptide synthesis: the adenylation (A) domain, the peptidyl carrier protein (PCP) and the condensation (C) domains. Peptides synthetized by NRPS are usually siderophores involved in iron uptake (e.g Vibriobactin from *Vibrio cholerae*) or have anti-microbial properties (e.g Bacitracin from *Bacillus licheniformis*) (Hur *et al.,* 2012). The AHL cluster and its cognate LuxR-type regulator (BRAD285_v2_0433-0436) as well as 3 NRPS clusters out of 4 are downregulated under bacteroid conditions (Lamouche *et al.,* 2018). However, the *ΔbclA* mutant displays a drastic upregulation of those NRPS clusters compared to WT strain at different intensities (Figure 5A). Although, synthetized peptide structure prediction is presumably inaccurate, hypotheses can be proposed about the probable role of these non-ribosomal peptides.

The first NRPS cluster could synthetize a putative Microsclerodermin (18% gene similarity), suggesting a defense role with a possible antifungal and/or antimicrobial activity (Hoffman *et al.,* 2013). It has also 27% gene similarity with a putative Ralsolamycin peptide. This peptide is involved in fungal hyphae colonization by *Ralstonia solanacearum*. This phytopathogen is able of endofungal lifestyle, which can be have important implications for its persistence in soils (Baldeweg *et al.,* 2017). This NRPS cluster seems to be involved in the free-living lifestyle of *Bradyrhizobium* sp. ORS285 and is correlated with the absence of *bclA* gene. Indeed, its expression in free living conditions is also upregulated in Δ*bclA* mutant compared to the WT. BclA could internalize and/or detect antimicrobial peptides at bacterial surface in free-living conditions, and its absence could constitutively activate the expression of this NRPS cluster which could synthetize a molecular weapon/signal leading to a better fitness in the rhizosphere. Distinct peptide uptake by BclA in the three conditions could induce this NRPS expression such as possible yeast extract derived Pathogen-associated Molecular Patterns (PAMPs) in the Yeast-Mannitol medium of the culture conditions or NCR-like peptides in bacteroid conditions.

The two other NRPS clusters do not display significant differences between the two genotypes in free living conditions, suggesting that their expression may not be a direct

consequence of *bclA* absence. The second NRPS cluster is responsible for the synthesis of a putative Nostopeptolide peptide (25% gene similarity). This cluster is downregulated in WT bacteroids compared to the free-living condition and upregulated in Δ*bclA* bacteroids compared to the WT ones. A functional symbiosis may then require the inhibition of this cluster expression. Moreover, the nostopeptolide appears to play a governing role during cellular differentiation symbiotic cyanobacterium *Nostoc punctiforme.* This metabolite inhibits the differentiation into hormogonium, which is the transient motile state of the cyanobacterium necessary for the initiation of a symbiotic interaction with plants. Nostopeptolide was found to be strictly down-regulated in symbiosis with *Gunnera manicata* and *Blasia pusilla* (Liaimer *et al.,* 2015). Although the real structure of this non-ribosomal peptide remains to be determined, the parallel between this putative metabolite produced by *Bradyrhizobium* sp. ORS285 and the nostopeptolide is very interesting. Indeed, in Δ*bclA* bacteroids, its presence can be a consequence of an upstream dysfunctional symbiotic signaling caused by *Aeschynomene* NCR-like peptides in the absence of BclA transporter.

The last NRPS cluster is only upregulated in free living and *Aeschynomene indica* (AI-Δ*bclA*) bacteroid conditions. The putative metabolite has no hit of gene similarity in ANTISMASH database. The AHL cluster is upregulated only in AI-Δ*bclA* condition. However, these hypotheses require to be nuanced by the absence of NRPS genes in *Bradyrhizobium* sp. ORS278, suggesting a probable exclusive involvement of non-ribosomal peptides in the free-living lifestyle of ORS285 (Giraud *et al.,* 2007). Finally, the upregulation of NRPS genes in the Δ*bclA* mutant could also be seen as an “aggressive” response, in which a bacterium with a reduced capacity to cope with AMPs (ie. absence of BclA) would produce antimicrobial compounds through NRPS to kill other bacteria before they produce those AMPs. Functional characterization of NRPS products is very challenging, but it might be a further step in the understanding of the biological role of this upregulation of NRPS genes in the Δ*bclA* mutant.

Modification of the regulation of photosynthesis and Calvin-Benson-Bassham cycle

Calvin-Benson-Bassham (CBB) cycle is upregulated in *bclA* bacteroids compared to the WT ones (Figure 5B). One of the two RubisCOs of *Bradyrhizobium* sp. ORS278 is required for the establishment of a functional symbiosis with *A. indica*. The suggested role of RubisCO in nitrogen-fixing symbiosis, and by extension of the CBB cycle, is to play a role of a transient electron sink.

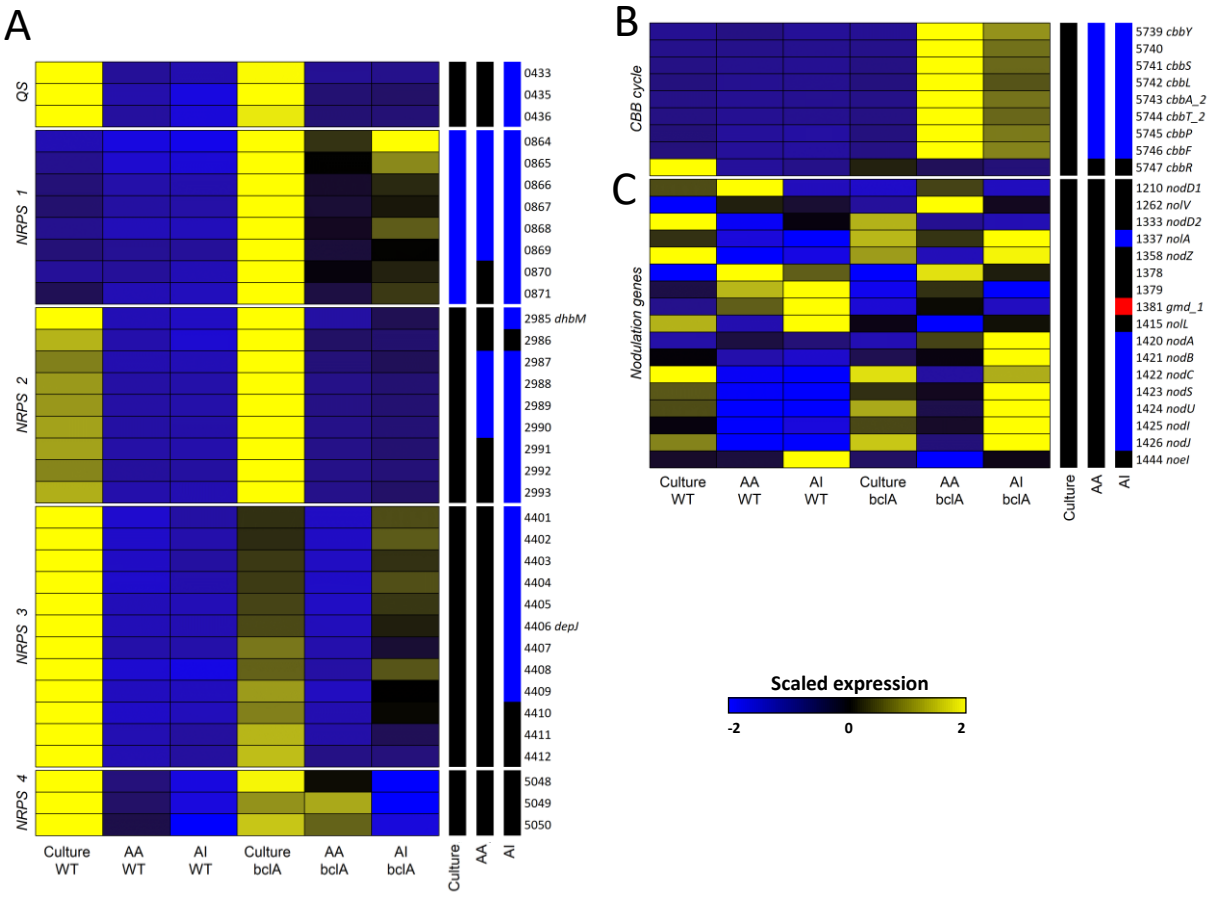


Figure 5: *bclA* mutation triggers the expression of Non-Ribosomal Peptide Synthases (NRPS), genes involved in the Calvin-Benson-Bassham cycle (CBB) and nodulation genes of *Bradyrhizobium* sp. ORS285 in bacteroids. Heatmaps showing the gene expression of NRPS clusters, CBB cycle and nodulation in *Bradyrhizobium* sp. ORS285 genome. NRPS are usually expressed in culture conditions, and Δ*bclA* bacteroids display a higher expression of 3 among these 4 clusters. Nodulation genes are induced only in *A. indica* nodules elicited by Δ*bclA* strain. Right columns indicate whether the gene is induced in Δ*bclA* (blue), WT condition (red) or not (black) in the three conditions (fdr<0.01 and |LFC|>1.58).

Indeed, a high level of reductive power is transferred to the bacteroids via dicarboxylate import and Krebs cycle activity. RubisCO can balance a redox critical state before the establishment of a functional nitrogenase activity, where O_2 concentration is not sufficient to balance the redox state of the cell by respiration (Gourion *et al.,* 2011). The same hypothesis could be used in *Bradyrhizobium* sp. ORS285 Δ*bclA* bacteroids insofar as the absence of nitrogenase activity cannot compensate the possible high input of carbon sources from the host plant, leading to the continuous expression of CBB cycle genes. This would also mean that the plant is not sanctioning the absence of effective nitrogen fixation by stopping carbon transfer to the bacterium, although sanctions have been shown to be applied to non-cooperative rhizobia in other legume rhizobium systems (Kiers *et al.*, 2003; Westhoek *et al.*, 2017).

bclA deletion induces higher expression of *nod* genes in *A. indica* bacteroids

A rather surprising result is the high upregulation of genes involved in Nod factor synthesis in *A. indica* bacteroids (Figure 5C), especially considering that *A. indica* and ORS285 strain establish a Nod factor independent symbiotic interaction (Giraud *et al.,* 2007; Bonaldi *et al.,* 2011). It suggests a probable absence of distinctive response from the bacterial strain to the different host plant flavonoids, leading to Nod factor synthesis even if it is not required to establish the symbiosis. Expression of Nod factor biosynthetic genes might then be repressed once a certain checkpoint in bacteroid differentiation is attained. This finally suggests that in case of *A. indica,* flavonoid content in nodules may be high as it keeps inducing the expression of Nod factor biosynthetic genes.

Conclusion

In this study, whole nodule metabolomics combined to bacterial transcriptomics were used to: i) characterize at the molecular level the dysfunctions that are encountered in the Δ*bclA* mutant in free living conditions as well as on symbiosis with its hosts *A. indica* and *A. afraspera*; ii) identify key mechanisms of bacteroid differentiation and nitrogen fixation by comparing the wild type bacteroid transcriptome to the one of Δ*bclA* mutant, used as an intermediate state between bacteria living outside the plant and functional differentiated bacteroids. We found that the Δ*bclA* mutation has limited impact on the bacterial transcriptome in free living conditions, whereas the *in planta* transcriptome was deeply altered as compared to the wild type. This observation is consistent with phenotypic analyses comparing WT and Δ*bclA* mutant strains in symbiosis with *A. afraspera* and *A. indica* (Guefrachi *et al.*, 2015; Barrière *et al.*, 2017).

The additional features brought by Δ*bclA* conditions allowed us to discriminate between bacterial genes that are commonly induced during nitrogen fixation and the ones that presumably respond to common environmental signals present in the nodules of the two distinct *Aeschynomene* species considered. In the same vein, Δ*bclA* bacteroid conditions also helped us to probably discriminate between host-specific genes that are induced and required for the nitrogen fixing host-specific interaction and the ones responding to not essential but host specific environmental signals BclA protein activity could be considered as a checkpoint between these distinct sets of genes which we integrated into a schematic model representing the different steps of intracellular accommodation of *Bradyrhizobium* sp. ORS285 in

Aeschynomene nodules (Figure 6). Finally, the higher expression of NRPS genes in Δ*bclA* mutant conditions suggests that BclA function could go beyond the establishment of the nitrogen fixing symbiosis and be involved in the regulation of defense mechanisms during free-living lifestyle of *Bradyrhizobium* sp. ORS285.

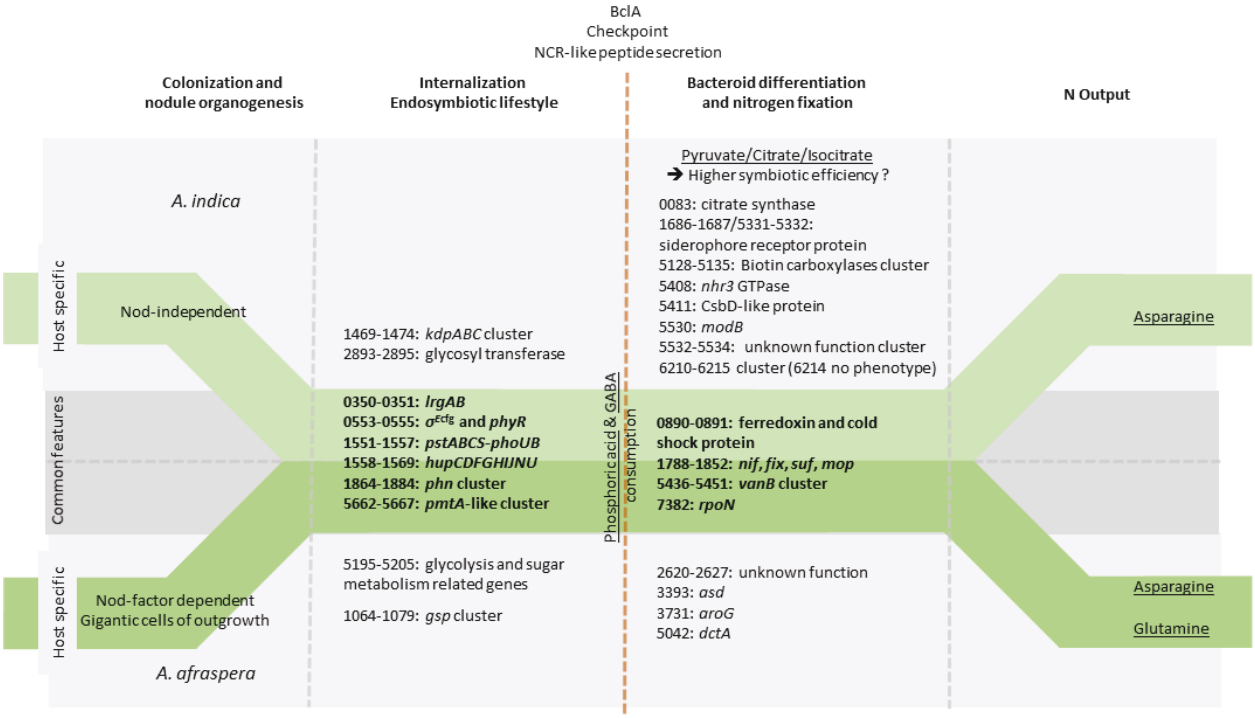


Figure 6: Model of *Bradyrhizobium* sp. ORS285 successive intracellular accommodation steps in *Aeschynomene* spp. nodules proposed from the transcriptomic and metabolomic analyses of Δ*bclA* bacteroid conditions. Bold genes are commonly induced while not bold ones are host-specific. Underlined words are relevant metabolites found in metabolomic analyses

Acknowledgements

F.L. was supported by a PhD fellowship from the Université Paris-Sud. This work was funded by the Agence Nationale de la Recherche, grant n° ANR-13-BSV7-0013 and used resources from the National Office for Research, Development and Innovation of Hungary, grant n° 120120 to A.K.

Supplemental material

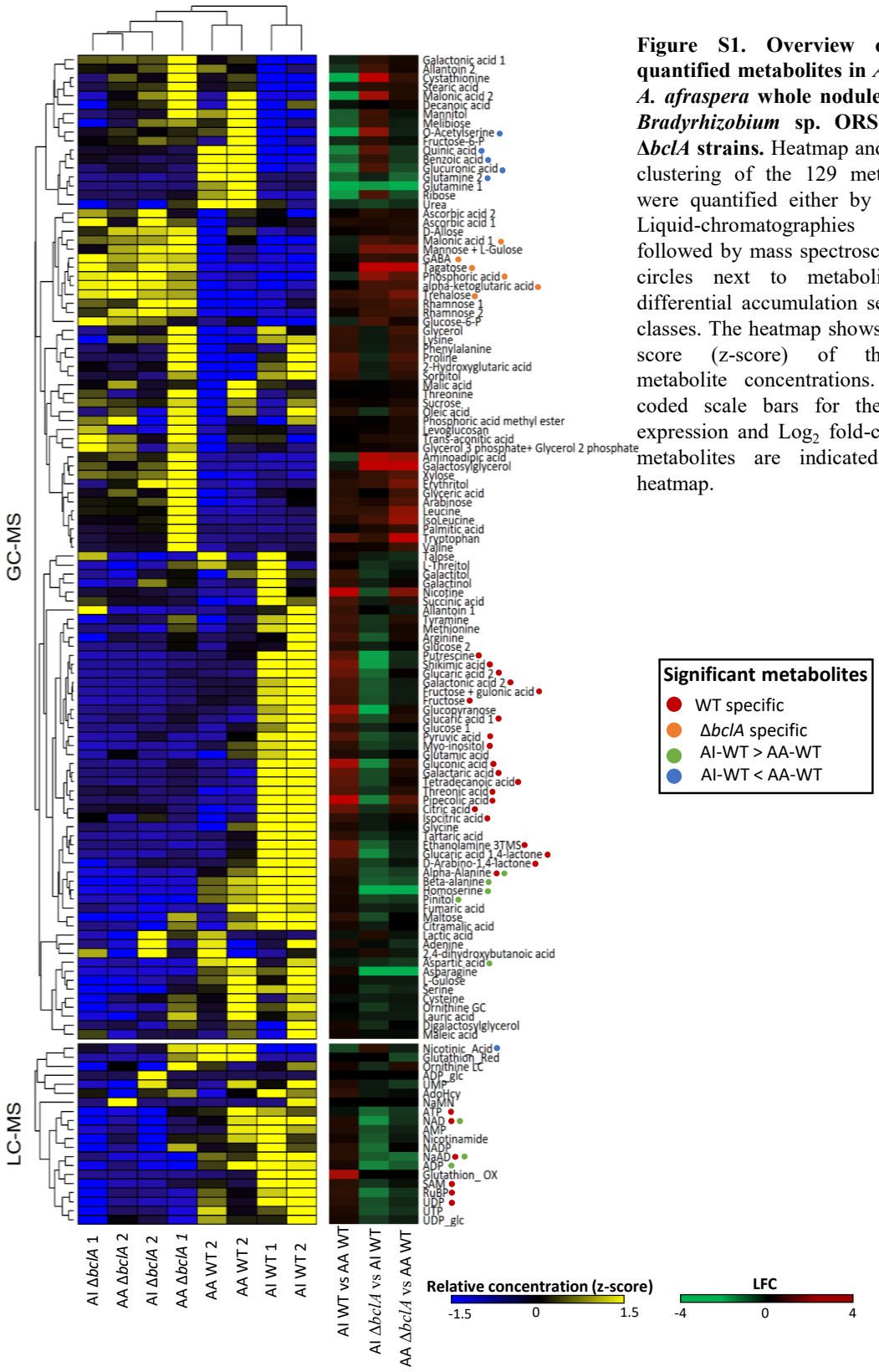


Figure S1. Overview of the 129 quantified metabolites in *A. indica* and *A. afraspera* whole nodules elicited by *Bradyrhizobium* sp. ORS285 WT or Δ*bclA* strains. Heatmap and hierarchical clustering of the 129 metabolites that were quantified either by (GC-MS) or Liquid-chromatographies (LC-MS) followed by mass spectroscopy. Colored circles next to metabolites indicate differential accumulation separated in 4 classes. The heatmap shows the standard score (z-score) of the obtained metabolite concentrations. The color-coded scale bars for the normalized expression and Log_2 fold-change of the metabolites are indicated below the heatmap.

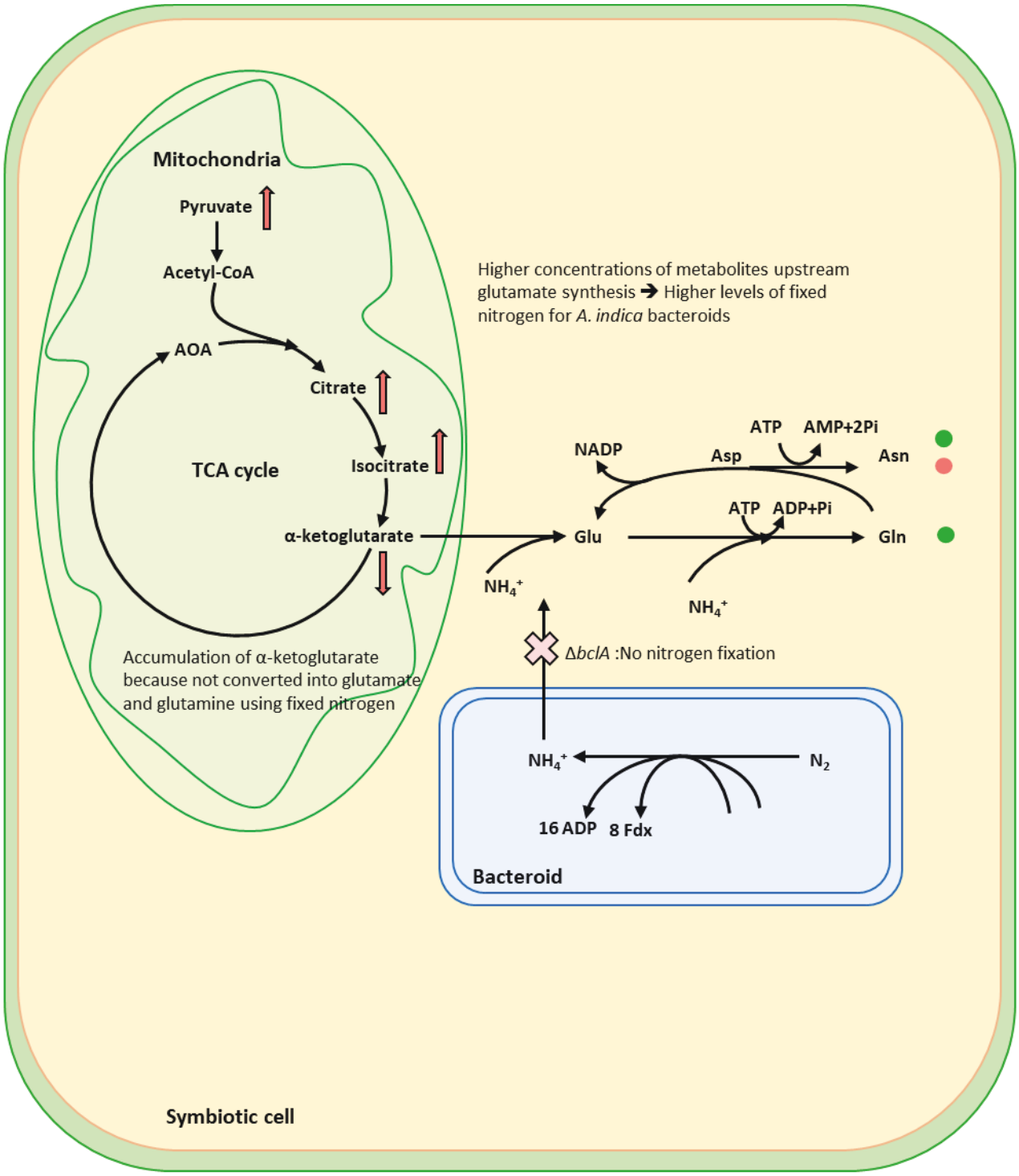


Figure S2. Metabolic model of the differential interplay between bacteroid ammonium secretion and mitochondrial metabolism depending on the plant host and bacteroid genotype. Pink arrows display differential accumulation of metabolites between *A. indica* and *A. afraspera* nodules elicited by *Bradyrhizobium* sp. ORS285. Pink and green circles show the output nitrogen metabolites for *A. indica* and *A. afraspera*, respectively.

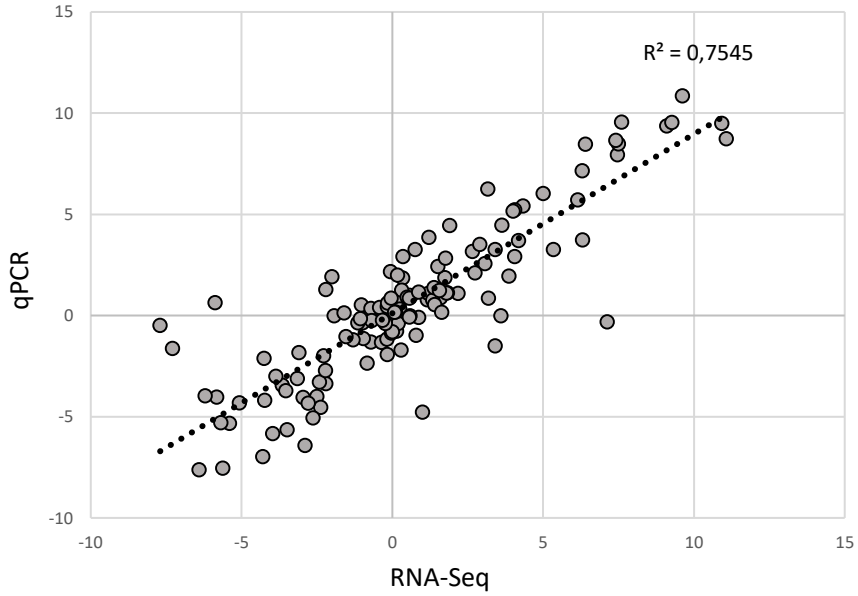
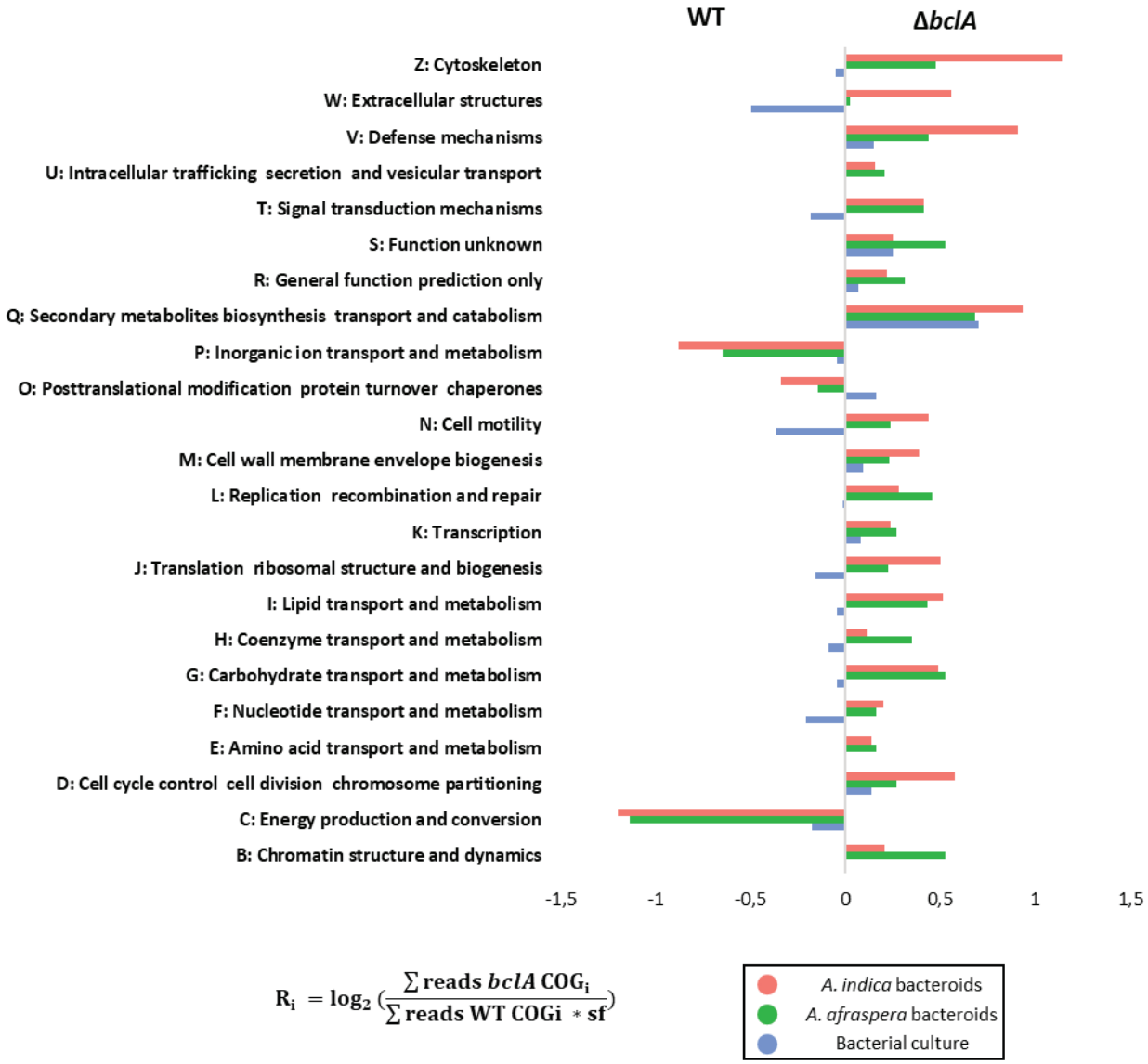


Figure S3. qRT-PCR cross validation of RNA-seq results. qPCR samples were obtained on three independent experiments, based on the log2-transformed fold change (LFC) expression values.



$$R_i = \log_2\left(\frac{\sum \text{reads } bclA \text{ COG}_i}{\sum \text{reads WT COGi} * sf}\right)$$

Figure S4. Read repartition ratio between Δ*bclA* mutant and WT of the reads among COG classes in the three comparisons. While COG classes C and P are lowered in Δ*bclA* conditions, classes D,Q and V and T are enriched.

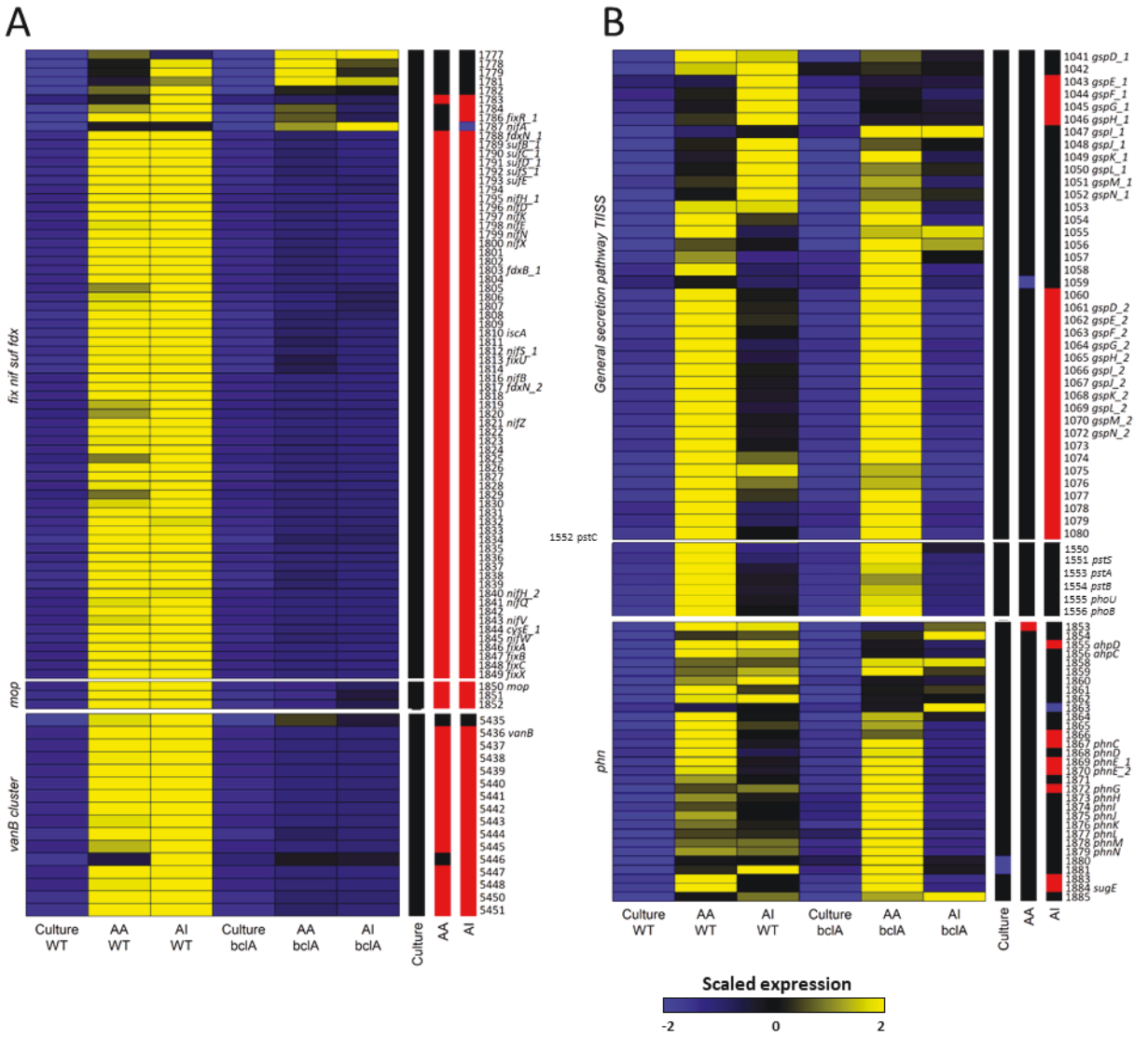


Figure S5: *bclA* mutation can either inhibit or have no effect on the expression of *Bradyrhizobium* sp. ORS285 bacteroid induced genes. Heatmaps showing the gene expression genes highly induced in wild-type bacteroid conditions (Lamouche et al., 2018). (A) The region containing genes involved in nitrogen fixation process, such as *nif, fix, suf, fdx* and *mop* genes, as well as *vanB* cluster previously identified (Lamouche et al., 2018) display drastic downregulation in Δ*bclA* bacteroids. (B) Other genomic clusters do not display such downregulations, such as the type II secretion system (T2SS), as well as *pst*/*pho* and *phn* cluster, involved in phosphate and phosphonate transport and metabolism. Right columns indicates whether the gene is induced in Δ*bclA* (blue), WT condition (red) or not (black) in the three conditions (fdr<0.01 and |LFC|>1.58).

References

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.

Arnold, M.F.F., Haag, A.F., Capewell, S., Boshoff, H.I., James, E.K., McDonald, R., *et al.* (2013) Partial complementation of *Sinorhizobium meliloti bacA* mutant phenotypes by the *Mycobacterium tuberculosis* BacA protein. *J Bacteriol* **195**: 389–98.

Baldeweg, F., Kage, H., Schieferdecker, S., Allen, C., Hoffmeister, D., and Nett, M. (2017) Structure of ralsolamycin, the interkingdom morphogen from the crop plant pathogen *Ralstonia solanacearum*. *Org Lett* **19**: 4868–4871.

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.

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.

Bonaldi, K., Gourion, B., Fardoux, J., Hannibal, L., Cartieaux, F., Boursot, M., *et al.* (2010) Large-scale transposon mutagenesis of photosynthetic Bradyrhizobium sp. strain ORS278 reveals new genetic loci putatively important for Nod-independent symbiosis with Aeschynomene indica. *Mol Plant-Microbe Interact* **23**: 760–770.

Bueno, E., Mesa, S., Sanchez, C., Bedmar, E.J., and Delgado, M.J. (2010) NifA is required for maximal expression of denitrification genes in *Bradyrhizobium japonicum*. **12**: 393–400.

Chapelle, E., Alunni, B., Malfatti, P., Solier, L., Pédron, J., Kraepiel, Y., and 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.

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., *et al.* (2015) Convergent evolution of rndosymbiont differentiation in Dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by nodule-specific cysteine-rich peptides. *Plant Physiol* **169**: 1254–1265.

Domenech, P., Kobayashi, H., LeVier, K., Walker, G.C., and Barry, C.E. (2009) BacA, an ABC transporter involved in maintenance of chronic murine infections with mycobacterium tuberculosis. *J Bacteriol* **191**: 477–485.

Dylan, T., Ielpi, L., Stanfield, S., Kashyap, L., Douglas, C., Yanofsky, M., *et al.* (1986) *Rhizobium meliloti* genes required for nodule development are related to chromosomal virulence genes in *Agrobacterium tumefaciens*. *Proc Natl Acad Sci U S A U S A* **83**: 4403–7.

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.-C., *et al.* (2007) Legumes symbioses: absence of nod genes in photosynthetic bradyrhizobia. *Science* **316**: 1307–1312.

Glazebrook, J., Ichige, A., and Walker, G.C. (1993) A *Rhizobium meliloti* homolog of the *Escherichia coli* peptide-antibiotic transport protein SbmA is essential for bacteroid development. *Genes Dev* **7**: 1485–1497.

Gourion, B., Delmotte, N., Bonaldi, K., Nouwen, N., Vorholt, J.A., and Giraud, E. (2011) Bacterial RuBisCO is required for efficient *Bradyrhizobium*/*Aeschynomene* symbiosis. *PLoS One* **6**: 1–9.

Gourion, B., Sulser, S., Frunzke, J., Francez-Charlot, A., Stiefel, P., Pessi, G., *et al.* (2009) The PhyR-σEcfGsignalling cascade is involved in stress response and symbiotic efficiency in Bradyrhizobium japonicum. *Mol Microbiol* **73**: 291–305.

Gu, Z., Eils, R., and Schlesner, M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. **32**: 2847–2849.

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.

Gully, D., Teulet, A., Busset, N., Nouwen, N., Fardoux, J., Rouy, Z., *et al.* (2017) Complete genome sequence of *Bradyrhizobium* sp. ORS285, a photosynthetic strain able to establish Nod factor-eependent or Nod factor-independent symbiosis with *Aeschynomene* legumes. *Genome Announc* **5**: e00421-17.

Haag, A.F., Baloban, M., Sani, M., Kerscher, B., Pierre, O., Farkas, A., *et al.* (2011) Protection of *Sinorhizobium* against host cysteine-rich antimicrobial peptides is critical for symbiosis. *PLoS Biol* **9**: e1001169.

Hoffmann, T., Müller, S., Nadmid, S., Garcia, R., and Müller, R. (2013) Microsclerodermins from terrestrial myxobacteria: An intriguing biosynthesis likely connected to a sponge symbiont. *J Am Chem Soc* **135**: 16904–16911.

Hur, G.H., Vickery, C.R., and Burkart, M.D. (2012) Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat Prod Rep* **29**: 1074–1098.

Kiers, E.T., Rousseau, R.A., West, S.A., and Denison, R.F. (2003) Host sanctions and the legume–rhizobium mutualism. *Nature* **425**: 78–81.

Lamouche, F., Gully, D., Chaumeret, A., Nouwen, N., Verly, C., Pierre, O., *et al.* (2018) Transcriptomic dissection of *Bradyrhizobium* sp. strain ORS285 in symbiosis with *Aeschynomene* spp. inducing different bacteroid morphotypes with contrasted symbiotic efficiency. *Environ Microbiol* doi:10.1111/1462-2920.14292:

Lê, S., Josse, J., and Husson, F. (2008) FactoMineR: An R package for multivariate analysis. *J Stat Softw* **25**: 1–18.

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.

LeVier, K., Phillips, R.W., Grippe, V.K., Roop, R.M., and Walker, G.C. (2000) Similar requirements of a plant symbiont and a mammalian pathogen for prolonged intracellular survival. *Science* **287**: 2492–3.

Love, M.I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 1–21.

Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., and Kondorosi, E. (2003) A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiol* **132**: 161–73.

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.

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.

Oldroyd, G.E.D., Murray, J.D., Poole, P.S., and Downie, J.A. (2011) The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev Genet* **45**: 119–144.

Prell, J., Bourdés, A., Karunakaran, R., Lopez-Gomez, M., and Poole, P. (2009) Pathway of γ-aminobutyrate metabolism in *rhizobium leguminosarum* 3841 and its role in symbiosis. *J Bacteriol* **191**: 2177–2186.

Stanfield, S.W., Ielpi, L., O'Brochta, D., Helinski, D.R., and Ditta, G.S. (1988) The *ndvA* gene product of *Rhizobium meliloti* is required for beta-(1-2)glucan production and has homology to the ATP-binding export protein HlyB. *J Bacteriol* **170**: 3523–3530.

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.

Sulieman, S. (2011) Does GABA increase the efficiency of symbiotic N_2 fixation in legumes? *Plant Signal Behav* **6**: 32–6.

Sulieman, S., Ha, C. Van, Schulze, J., and Tran, L.S.P. (2013) Growth and nodulation of symbiotic *Medicago truncatula* at different levels of phosphorus availability. *J Exp Bot* **64**: 2701–2712.

Sulieman, S. and Schulze, J. (2010) Phloem-derived γ-aminobutyric acid (GABA) is involved in upregulating nodule N2 fixation efficiency in the model legume Medicago truncatula. *Plant, Cell Environ* **33**: 2162–2172.

Tiricz, H., Szücs, 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.

Udvardi, M. and Poole, P.S. (2013) Transport and Metabolism in Legume-Rhizobia Symbioses. *Annu Rev Plant Biol* **64**: 781–805.

Vallenet, D., Calteau, A., Cruveiller, S., Gachet, M., Lajus, A., Josso, A., *et al.* (2017) MicroScope in 2017: An expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res* **45**: D517–D528.

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.

Wang, D., Griffitts, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., *et al.* (2010) A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* **327**: 1126–1129.

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Bruccoleri, R., *et al.* (2015) AntiSMASH 3.0-A comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* **43**: W237–W243.

Wehmeier, S., Arnold, M.F.F., Marlow, V.L., Aouida, M., Myka, K.K., Fletcher, V., *et al.* (2010) Internalization of a thiazole-modified peptide in *Sinorhizobium meliloti* occurs by BacA-dependent and -independent mechanisms. **156**: 2702–2713.

Westhoek, A., Field, E., Rehling, F., Mulley, G., Webb, I., Philip, S., *et al.* (2017) Policing the legume-Rhizobium symbiosis: a critical test of partner choice. *Sci Rep* **7**: 1419.

White, J., Prell, J., James, E.K., and Poole, P. (2007) Nutrient Sharing between Symbionts. *Plant Physiol* **144**: 604–614.

4. Atypical terminal differentiation of *Bradyrhizobium diazoefficiens* USDA110 bacteroid in *Aeschynomene afraspera* analyzed through multi-omics approach

Florian Lamouche, Quentin Nicoud, Anaïs Chaumeret, Thierry Baliau, Florence Guérard, Erika Sallet, Solenn Tuffigo, Olivier Pierre, Yves Dessaux, Francoise Gilard, Bertrand Gakière, Istvan Nagy, Attila Kereszt, Michel Zivy, Peter Mergaert, Benjamin Gourion et Benoit Alunni

Avant-propos

Le but de cette étude était d'étudier les signatures moléculaires de la différenciation des bactéroïdes liées à l'effet des peptides NCR-like. Pour ce faire, il est nécessaire de comparer la physiologie d'une souche bactérienne capable de noduler une plante hôte n'induisant pas de différenciation des bactéroïdes et une plante induisant cette différenciation. Il se trouve que la plante hôte *A. afraspera* peut également être nodulée par *B. diazoefficiens* USDA110 (Renier *et al.*, 2011). Cette souche est un symbionte naturel du soja qui n'impose pas de programme de différenciation des bactéroïdes. Nous avons donc comparé les transcriptomes et les protéomes de la souche USDA110 en symbiose avec *A. afraspera* et le soja. Nous avons également effectué un métabolome de nodosités complètes et des analyses fonctionnelles de gènes d'intérêt ressortant de cette intégration de jeux de données-omiques. Cette étude comprend aussi une comparaison de l'efficacité symbiotique et de la survie des bactéroïdes d'USDA110 entre *A. afraspera* et *G. max*.

Ce travail a également débuté avant mon arrivée en stage de M2, par la production du jeu de données RNA-seq comprenant une culture de référence en milieu riche (YM) et des bactéroïdes de soja et de *A. afraspera* récoltés à 14dpi. J'ai généré pendant mon M2 le protéome pour les mêmes conditions. J'ai ensuite analysé les résultats de RNA-seq comme pour les autres chapitres, et intégré à ces derniers les résultats de protéomique en collaboration avec la plateforme protéomique PAPPSO. J'ai également analysé les résultats de métabolomique en nodosités entières. J'ai effectué des analyses fonctionnelles basées sur ces résultats dont 9 mutants d'insertion et 4 de délétions que j'ai généré avec l'aide de Solenn Tuffigo (stagiaire M1). J'ai aussi testé d'autres mutants déjà existants dans la littérature et dont j'ai testé le phénotype de tous ces mutants *in planta*. J'ai également mis au point et participé aux expériences quantifiant la survie et la perméabilité des bactéroïdes avec Quentin Nicoud.

Abstract:

Legume plants can interact with soil bacteria, referred to as rhizobia, and house them within a dedicated organ called nodule. There, the intracellular rhizobia, the bacteroids, fix nitrogen for the benefit of the plant. The specificity of rhizobium-legume associations is, in general, very high and a given rhizobial species is adapted to a restricted number of legume hosts. In natural environment, amongst the couples able to form nodules, the whole spectrum of efficiency varying from non-fixing associations to interactions that are highly beneficial for the plant, can be encountered. *Bradyrhizobium diazoefficiens* USDA110 can establish a functional symbiotic interaction with soybean and with *Aeschynomene afraspera*. In contrast to soybean, *A. afraspera* was previously reported to trigger the terminal differentiation of its bacteroids but typical features of this differentiation were surprisingly not observed upon nodulation with USDA110. Herein we show that USDA110 and *A. afraspera* establish a poorly efficient symbiosis during which USDA110 bacteroids undergo an atypical terminal differentiation. Combined omics approaches were used to study the physiology of USDA110 bacteroids under terminally and non-terminally differentiated states using two host plants. Strong differences in gene expression/protein accumulation between the two host environments were observed suggesting drastic and host-specific modifications of bacteroid physiology, and especially the high level of stress of *A. afraspera* bacteroids. This study pinpoints to which extent a given rhizobium can adapt its physiology to a new host and how rhizobia cope with terminal differentiation when they did not co-evolve with such hosts.

Introduction:

Nitrogen availability is a major limitation for the development of plants in many environments including agricultural settings. To overcome nitrogen deprivation and to thrive on substrates presenting a low nitrogen content, plants are heavily fertilized, causing important environmental concerns and financial drawbacks (Erisman *et al.*, 2013, Zhao *et al.*, 2017). Being able to form symbiotic associations with soil bacteria, the rhizobia, that fix atmospheric nitrogen for their benefit, plants of the legume family display a critical asset for their development on such a nitrogen poor substrate. On the root, those associations lead to the development of rhizobia housing organs called nodules. In plants, the rhizobia adopt an intracellular lifestyle and differentiate within plasmalemma-derived vesicles. Under this form, referred to as bacteroids, rhizobia convert atmospheric dinitrogen into ammonium conferring to the plant a decisive advantage to thrive on nitrogen scarce substrates.

Not all the rhizobium-legume interactions result in a benefice for the plants. A relatively high level of specificity exists in these interactions where one rhizobial genus, species, or strain will usually be only capable to reach the nitrogen fixing state with a restricted number of genus, species or ecotypes of the legume clade. Critical recognition steps occur all along the symbiotic process (Wang *et al.*, 2017; Yang *et al.*, 2017; Gourion and Alunni, 2018; Stonoha-Arther and Wang, 2018). 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 impact not only the tolerance to rhizobia but also the efficiency of the symbiosis.

Recently, Nodule specific Cysteine-Rich (NCR) peptides produced by some, but not all legumes, were proposed to play a crucial role in the control of host-symbiont specificity at the intracellular stage of the symbiosis (Gourion and Alunni, 2018). Those peptides are targeted to the bacteroid containing compartments where they govern the bacteroid differentiation in legumes of the Inverted Repeat Lacking Clade (IRLC) (Wojciechowski *et al.*, 2000; Mergaert *et al.*, 2003; Alunni *et al.*, 2007; Van de Velde *et al.*, 2010; Guefrachi *et al.*, 2014). In this clade, the differentiation processes are so important that they suppress the bacteroid capacity to resume growth. Therefore, they are referred to as terminal bacteroid differentiation (TBD). Systematically, TBDs have been associated with changes in bacteroid shape (cell elongation) and an increase in bacteroid DNA content through a cell cycle switch toward endoreduplication (Mergaert *et al.*, 2006; Czernic *et al.*, 2015; Montiel *et al.*, 2017). Furthermore, an increased permeability of the bacteroid envelope also occurs in TBD triggering legumes, most probably

due to the interaction of NCR peptides with bacterial membranes (Mergaert *et al.*, 2006; Alunni and Gourion, 2016). Similarly, NCR like peptides have been identified in the phylogenetically distant genus *Aeschynomene*, where they are also associated to TBD resulting in either elongated or spherical polyploidy bacteroids (Czernic *et al.*, 2015).

Interestingly some peculiar rhizobia can nodulate a large array of plant species. One of them, *Bradyrhizobium diazoefficiens* strain USDA110, formerly *B. japonicum*, can trigger functional nodules without TBD on soybean (*Glycine max*), cowpea or siratro (Koch *et al.*, 2010). In addition to these species, USDA110 nodulates *Aeschynomene afraspera*, a distant plant species from the above-mentioned ones in terms of phylogeny, physiology and morphology (Figure 1A; Renier *et al.*, 2011; Ledermann *et al.*, 2018). Surprisingly, in *A. afraspera*, USDA110 shows only very limited features that have been systematically associated with TBD suggesting that the bacterium might be resistant to the TBD process (Barrière *et al.*, 2017). To our knowledge such resistance against TBD was only reported for another symbiotic couple in the literature (Crespo-Rivas *et al.*, 2016; Alunni and Gourion, 2016; Barrière *et al.*, 2017) and the physiology of the corresponding bacteroids was not analyzed at a global scale.

Herein, we further characterize the symbiosis established between USDA110 and *A. afraspera* with special emphasis on bacteroid differentiation. Our observations, supported by whole nodule metabolome analysis, indicate that USDA110 is poorly matched for nitrogen fixation with *A. afraspera*. To better understand the adaptation of *B. diazoefficiens* on *G. max* and on *A. afraspera*, we used a combination of transcriptomics (RNA-seq) and shotgun proteomics (LC-MS/MS) approaches that provide complementary information. Such combination of omics approaches was very rarely used to explore the physiology of legume bacteroids and was never employed to study adaptation of bacteroids to such distant hosts differing in their capacity to produce NCRs and to induce bacteroid differentiation (Delmotte *et al.*, 2010; Koch *et al.*, 2010; diCenzo *et al.*, 2018). Finally, despite previously published observations suggesting that USDA110 is insensitive to NCRs (Kulkarni *et al.*, 2015; Barrière *et al.*, 2017), we report here that USDA110 undergoes a terminal and very atypical bacteroid differentiation in *A. afraspera*.

Material and Methods:

Bacterial cultures and bacteroid extraction

Bradyrhizobium diazoefficiens strain USDA110 (Regensburger and Hennecke, 1983) was cultivated in Yeast Mannitol medium at 30°C in a rotary shaker (Giraud *et al.*, 2000). Exponential culture samples were grown until OD_{600nm} reached 0,5. For transcriptomic analysis only, samples were transferred on ice for 10 minutes in presence of 10% (v/v) of an ice-cold phenol/ethanol mixture (5:95, v/v). Subsequently, bacteria were collected by centrifugation (10 min, 3500g) and frozen in liquid nitrogen prior to -80°C storage. *G. max* ecotype Williams 82 (Schmutz *et al.*, 2010) and *A. afraspera* seeds were surface sterilized and the plants were cultivated as described in Barrière *et al.*, (2017). Nodules were collected at 14 days post inoculation (dpi) and immediately immersed in liquid nitrogen and stored at -80°C until use. Each tested condition (*in vitro* and *in planta*) has been produced in three biological replicates.

Analysis of TBD features

Bradyrhizobium sp. ORS285 and *B. diazoefficiens* USDA110 strains used here were harboring a green/yellow fluorescent protein (Bonaldi *et al.*, 2011; Ledermann *et al.*, 2015) to better distinguish bacteria from electronic noise of flow cytometers. Bacteroids were extracted from freshly collected 14 dpi nodules according to Guefrachi *et al.*, (2015) and processed by flow cytometry using a CytoFLEX S (Beckman-Coulter). For ploidy measurements analyses, samples were heat-killed to be fully permeable to the DNA stain propidium iodide (PI, 50 µg. mL^{-1}). Preliminary permeability assays have been performed without heat-killing bacteria and their PI permeability was assessed over time. For bacteroid viability assays, we measured the bacterial concentration of the samples as well as the DNA content of single cells. Fifty microliters of bacterial suspension were serially diluted and plated on selective medium. Colony forming units (cfu) were counted 5 days post plating and allowed the determination of bacteroid viability. *Bradyrhizobium* sp. ORS285 in culture and in symbiosis with *Aeschynomene afraspera* were used as references for the analysis of TBD.

Genome annotation and RNA-seq analysis

Nodule and bacterial culture total RNA was extracted and treated as described in Lamouche *et al.* (2018). Oriented (strand-specific) libraries were produced using the SOLiD Total RNA-seq kit (Life Technologies) according to the manufacturer's instructions and were sequenced on a 40 M reads scale on a SOLiD 3 station to provide 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, and the mapping was performed using this new version of the genome (Sallet *et al.*, 2014). Quantitative analysis of the transcriptome was performed using DE-seq2 as described in Lamouche *et al.*, (2018). Genes were considered differentially expressed when they showed a $\log_2$ fold change (LFC) > 1,58 (ie. fold change >3) with a false discovery rate (FDR) < 0,01.

Proteomic analysis

Bacteroid were extracted from 14 dpi frozen nodules and the bacterial pellets were resuspended and lysed in -20°C acetone and sonicated (Mergaert *et al.*, 2006). Solubilization, dosage, digestion (trypsin 2% w/w) and solid phase extraction (using Phenomenex polymeric C18 column) were performed according to (Langella *et al.*, 2013). Peptides (800ng of proteins) were analyzed by LC-MS/MS with a Q Exactive mass spectrometer (Thermo Electron) using a nanoelectrospray interface (non-coated capillary probe, 10 µ i.d.; New Objective). Peptide ions were analyzed using Xcalibur 2.1 software. Proteins were quantified using spectral counting method (Delmotte *et al.*, 2014).

Metabolomic analysis

Fourteen dpi nodules were collected in liquid nitrogen and lyophilized. Metabolites and cofactors were extracted and analyzed by GC-MS and LC-MS respectively according to Su *et al.*, (2016).

Generation and phenotyping of bacterial 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 *Bradyrhizobium diazoefficiens* USDA110 and mutants were selected on 50 µg/mL kanamycin and 25 µg/mL cefotaxim plates. For the construction of deletion mutants, flanking regions of the gene were PCR amplified, fused by overlap extension PCR and cloned in a modified pNPTS129 suicide plasmid harboring YFP gene from Lederman *et al.*, (2015), yielding the plasmid pFLY. The resulting plasmids were introduced into *Bradyrhizobium diazoefficiens* USDA110 and mutants were obtained after a double recombination event, spotted by antibiotic resistance and fluorescence loss. 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).

Results

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

Previous reports indicate *B. diazoefficiens* USDA110, the model symbiont of *G. max*, is able to establish a functional nitrogen-fixing symbiosis with *A. afraspera*, a phylogenetically distant host belonging to the Dalbergioid clade that naturally interacts with photosynthetic rhizobia such as *Bradyrhizobium* sp. ORS285 (Figure 1A-C) (Renier *et al.*, 2011; Ledermann *et al.*, 2018; Barrière *et al.*, 2017). To evaluate the efficiency of this symbiosis, nitrogenase activity of USDA110- and ORS285-nodulated plants and their nitrogen content were determined by acetylene reduction assays and by elemental analysis respectively. Although nitrogenase activity was clearly detected in both types of nodules, it was significantly lower in USDA110-nodulated plants (Figure 1D). The same result is observed for mass gain per g of nodule, used in a previous study as a proxy of symbiotic efficiency measurements (Figure 1E; Lamouche *et al.*, 2018). Nitrogen and carbon contents were also reduced in USDA110-nodulated plants reaching levels of non-inoculated plants (Figure S1).

Moreover, their shoot/root ratio, a metrics that reflect the nutritional status of the plant, is reduced in USDA110-nodulated plants, indicating that the plant nutritional needs are not fulfilled and that the plant needs to expand its roots to find nutrients in the substrate (Figure S2). To characterize further this apparent dysfunction of the symbiosis, we analyzed the whole nodule metabolome, which mostly reflects the plant metabolism (Figure S3). We find that asparagine/glutamine may be the major type of exported nitrogen in *A. afraspera*, but the amount of the latter ones is reduced in USDA110-nodulated *A. afraspera* plants as compared to ORS285-nodulated plants. In addition, we find specifically in *A. afraspera*/USDA110 nodules the accumulation of sucrose, phosphoric acid and ascorbate, and oppositely, a strong reduction in trehalose content (Figure 1F).

In a nodule that functions to its maximal level, sucrose derived from phloem sap should be consumed to fuel the bacteroids with carbon substrates (usually dicarboxylates). Also, phosphate is often seen as a limiting factor in the nitrogen fixation process and the accumulation of phosphoric acid in nodules suggests that nitrogen fixation is not reaching its optimal rate (Hernandez *et al.*, 2009). Ascorbate has been shown to increase nitrogen fixation activity by modulating the redox status of leghemoglobin (Ross *et al.*, 1999; Bashor and Dalton, 1999).

Thus, its accumulation in nodules with reduced nitrogen fixation capacity can be seen as a strategy to rescue nitrogen fixation in an organ that does not fix nitrogen very efficiently. In agreement with these observations and the idea that functioning of USDA110-induced nodules is suboptimal, ORS285-nodulated *A. afraspera* plants display darker green leaves. Together these data indicate that USDA110 is poorly matched with *A. afraspera* for nitrogen fixation and suggest a metabolic disorder in this symbiotic interaction.

Overview of the USDA110 bacteroid proteomes and transcriptomes.

In order to better understand the misfit interaction between USDA110 and *A. afraspera*, the bacteroid physiology was analyzed through a transcripto-proteomic approach. The *G. max* was used as reference system in which USDA110 bacteroids are efficient and not exposed to NCRs. An additional reference, USDA110 cells cultivated in rich medium (exponential growth phase in aerobic condition) was also analyzed as an additional reference (Figure 2A).

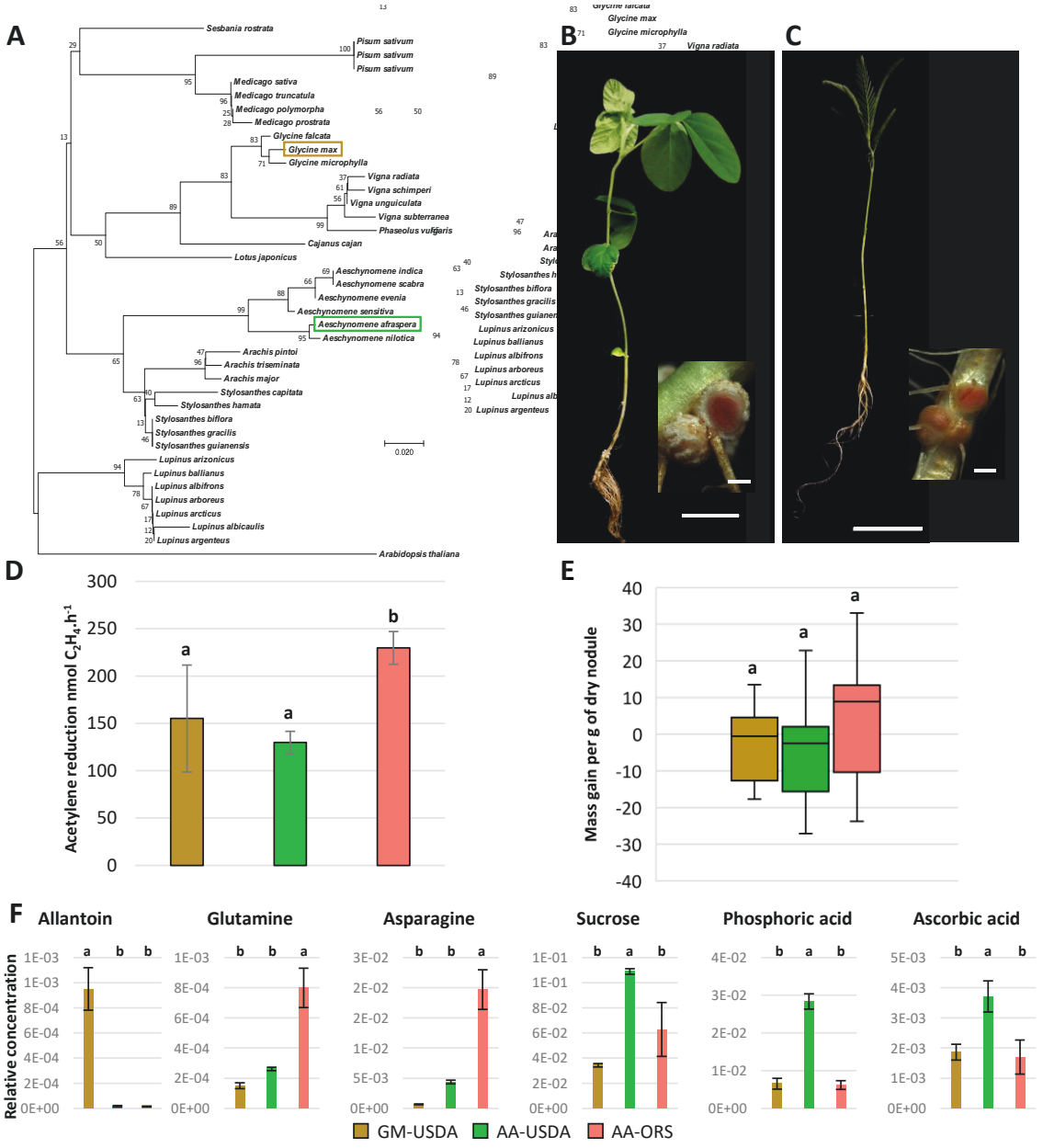


Figure 1. The non-adapted symbiotic couple formed by *Bradyrhizobium diazoefficiens* USDA110 and *Aeschynomene afraspera* displays suboptimal nitrogen fixation and nodule metabolic dysfunction. A, B. General aspect and nodule sections (inlay) of *G. max* (A) and *A. afraspera* (B) at 14dpi. Scale bars: 5cm (plants) and 1mm (nodules). C. Distance (Neighbour Joining method) tree of a selection of plant species from the Papilionoideae family based on *matK* sequences. D, E. Nitrogen fixation (D) and biomass of 14dpi plants (E). Total dry masses were measured and nitrogenase activity was determined by acetylene reduction assay. F. Whole nodule metabolome determined by GC/MS or LC/MS at 14dpi. 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 : *Glycine max*, AA : *Aeschynomene afraspera*, USDA / ORS : Nodules elicited by *B. diazoefficiens* USDA110 / *Bradyrhizobium* sp. ORS285.

The transcriptomic dataset was used to reannotate USDA110 genome using the EugenePP software (Sallet *et al.*, 2014), before quantification of transcript abundance and identification/quantification of protein accumulation. 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 USDA110 genome to 9171. Moreover, we also identified 246 ncRNA, ranging from 49 to 765 bp (median = 76 bp). 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 Supp Table 1.

Regarding the proteomic dataset, 1850 proteins have been identified. Principal component analysis (PCA) of each dataset revealed that samples are well partitioned by condition. For the proteomics dataset, 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 (15,1% of the observed variance; Figure 2B). A similar distribution of the samples on the PCA plot is obtained using the transcriptomic dataset, with a first axis explaining 42,6% of the observed variance and a second axis explaining 23,5% of the observed variance (Figure 2C).

Although differences are less pronounced in the proteomic dataset than in the transcriptomic dataset, COG analysis shows similar profiles across functional categories, except for membrane proteins that are less well identified in proteomics than transcriptomics (Figure S4). Among the 1850 proteins identified, 825 show differential accumulation and 443 of them are also differentially expressed in transcriptomic datasets (Figure 2D).

Moreover, we analyzed the Pearson correlation between transcriptomic and proteomic profiles and found that ~65% 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 ($r<-0,9$; Figure 2E-F). This observation suggests that the transcriptome (which provides a more exhaustive view than the proteome) and the proteome show globally a similar picture of bacterial physiology. Thus, our description of the bacterial functions will be primarily based on proteomes, and transcriptome will be used only when proteomics is not informative, for example to study known regulons and stimulons.

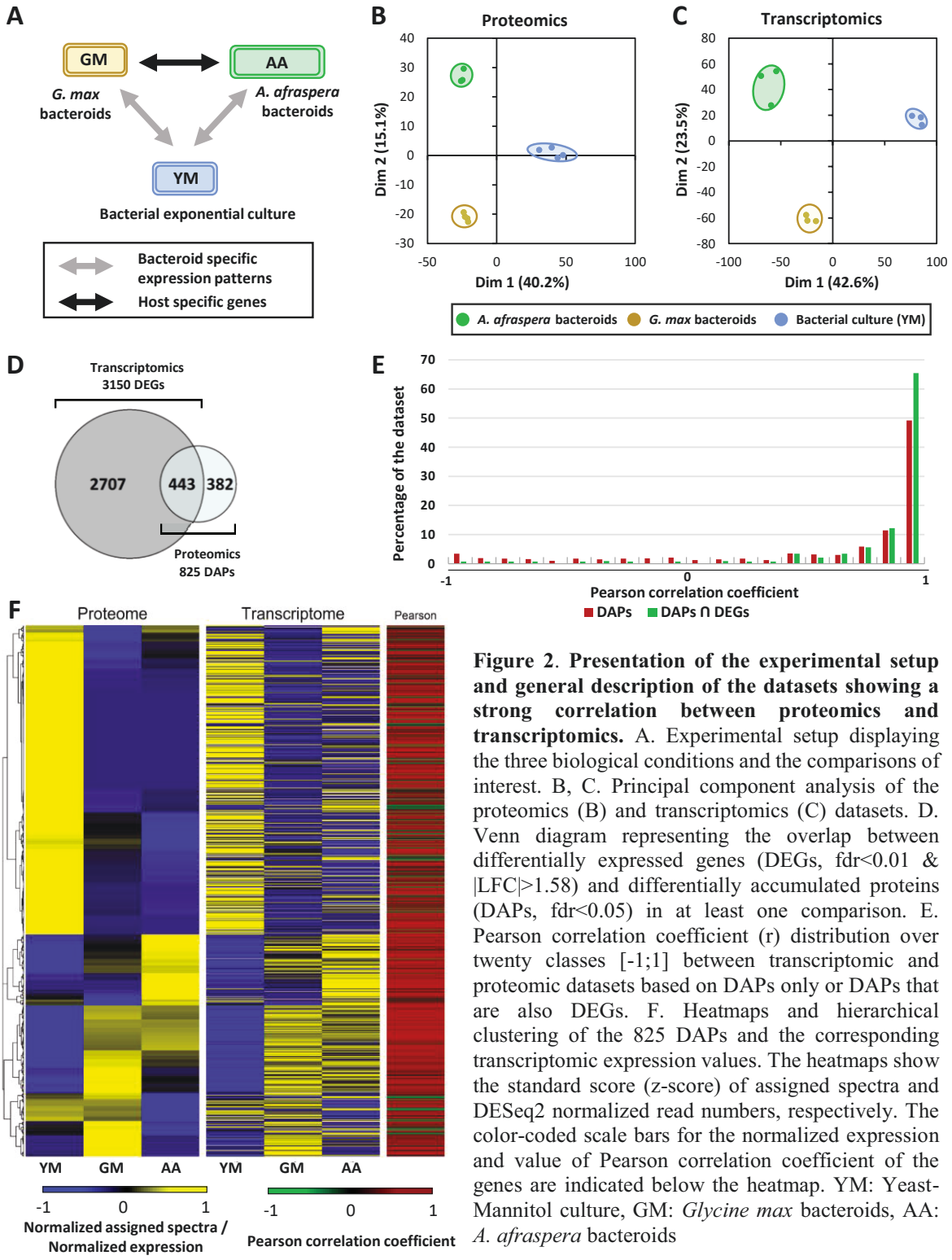


Figure 2. Presentation of the experimental setup and general description of the datasets showing a strong correlation between proteomics and transcriptomics. A. Experimental setup displaying the three biological conditions and the comparisons of interest. B, C. Principal component analysis of the proteomics (B) and transcriptomics (C) datasets. D. 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. E. Pearson correlation coefficient (r) distribution over twenty classes [-1;1] between transcriptomic and proteomic datasets based on DAPs only or DAPs that are also DEGs. F. Heatmaps and hierarchical clustering of the 825 DAPs and the corresponding transcriptomic expression values. The heatmaps show the standard score (z-score) of assigned spectra and DESeq2 normalized read numbers, 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: *Glycine max* bacteroids, AA: *A. afraspera* bacteroids

Symbiotic functions common to both types of USDA110 bacteroids

Statistical analysis performed on proteomics revealed that amongst the 1850 proteins detected during this study 825 showed significant variations between conditions observed by the ANOVA test. Among those, 712 and 700 proteins are significantly differentially accumulated in *G. max* and *A. afraspera* respectively as compared to the *in vitro* reference bacteria cultivated in rich medium. Strikingly, 645 proteins differentially accumulated in the *G. max*/*in vitro* comparison are also found to be differentially accumulated in the *A. afraspera*/*in vitro* comparison (Supplementary table 1).

Proteomics data provided a good view of the central metabolism of *B. diazoefficiens* upon exponential growth in aerobic nutrient-rich condition and in the nodules of the two hosts used in this study. Those data suggest that the central metabolism of bacteria residing in *A. afraspera* and *G. max* nodules is characterized by microoxic respiration and nitrogen fixation. Key enzymes involved in these two processes have been detected amongst the proteins harboring the highest spectra number somehow reflecting their abundance in the nodule samples (Table 1). This includes for instance, the nitrogenase and the nitrogenase reductase subunits. Those enzymes constitute the nitrogenase complex responsible for nitrogen conversion into ammonium. Similarly, FixA and FixB, involved in microaerobic respiration display two of the highest spectra number for nodule proteins while they were strictly not detected upon free-living condition.

In addition to these classical bacteroid functions, other proteins strongly accumulated in nodules were found. This is the case of 3 chaperonins (GroEL1, 5 and 6) Another example is the hydrogenase uptake large subunit which 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 ACC deaminase, blr0241, which was also amongst the most strongly accumulated proteins in nodules and was very significantly less abundant in free living USDA110.

In the transcriptomic dataset, 1999 differentially expressed genes (DEG) representing ~21% of the genome could be identified between the bacterial culture and the nodule environments regardless of the host. Among them, 1076 genomic objects displayed higher expression in nodules (including 7 newly annotated ncRNA and 1 newly annotated CDS among the 20 differentially expressed genes with highest fold change) and 923 genomic objects were

repressed *in planta* (including 2 newly annotated ncRNA and 2 newly annotated CDS among the 20 DEG with highest fold change, Supplementary Table 1).

Label	gene name	EugenePP annotation	mean SC AA	mean SC GM	mean SC YM	gene mutation
Nitrogen fixation and symbiotic islands (among 76 functions that are bacteroid induced)						
blr1743	*nifD*	IPR010143:Nitrogenase component 1, alpha chain	24[a]	24,3[a]	0,5[b]	
blr1744	*nifK*	IPR005976:Nitrogenase molybdenum-iron protein beta chain	56,7[a]	42,8[b]	3,3[c]	
blr1769	*nifH*	IPR000392:Nitrogenase iron protein, subunit NifH/Protochlorophyllide reductase, subunit ChlL	87,3[a]	78[b]	32[c]	
blr2037	*nifA*	IPR010113:Nif-specific regulatory protein	10[a]	6[b]	0[c]	
blr2038	*fixA*	IPR012255:Electron transfer flavoprotein, beta subunit	24,7[a]	21,3[b]	0[c]	
blr1773	*fixB*	IPR001308:Electron transfer flavoprotein, alpha subunit	20,7[a]	17,3[b]	0[c]	
Chaperonins						
bll2059	*groEL_1*	IPR002423:Chaperonin Cpn60/TCP-1 family	279,3[a]	250,8[b]	80,5[c]	
blr5626	*groEL_5*	IPR002423:Chaperonin Cpn60/TCP-1 family	188,3[a]	153,3[b]	145,5[c]	
blr6979	*groEL_6*	IPR002423:Chaperonin Cpn60/TCP-1 family	72,7[a]	65,3[b]	55,5[c]	
ACC synthase						
blr0241	blr0241	IPR001926:Tryptophan synthase beta subunit-like PLP-dependent enzyme	36,3[a]	32,3[b]	4[c]	
ABC transporters, solute binding proteins						
bll7921	bll7921	IPR000914:Solute-binding protein family 5 domain	21[b]	24,3[a]	0,5[c]	Yes
blr7922	blr7922	IPR000914:Solute-binding protein family 5 domain	15,7[a]	12[b]	0[c]	Yes
Lipid synthesis						
blr2148	*fpps_1*	IPR017446:Polyprenyl synthetase-related	4,3[a]	4[a]	0[b]	Yes

Table 1. Selection of *B. diazoefficiens* USDA110 proteins over-accumulated in bacteroids of both hosts as compared to the culture reference. Mean SC are the mean counted spectra among 4 replicates. Superscript letters represent significant differences after ANOVA and post hoc Tukey tests ($p<0.05$).

To go further, 4 loss of function mutants have been created in genes that displayed similar expression in both hosts, *fpps_1*, involved in terpene synthesis and an ABC transporter complex formed by two solute-binding proteins blr7922 and blr7921. We also deleted the permease and ATP-binding domain of this complex, blr7920-7918. Above all the tested mutants, phenotypic data has only been gathered ono the deletion mutants and one insertion mutant on the similar gene (blr7922), but all mutants quoted have been tested. However, none of them displayed any symbiotic phenotype (Figure S5).

Nevertheless, transcriptomic dataset also unveiled a gene cluster involved in phospholipid synthesis. This same cluster is also upregulated in *Bradyrhizobium* sp. ORS285 in bacteroid conditions (Lamouche *et al.*, 2018). We tested the phenotype of available mutants involved in phosphatidylcholine synthesis, encoding phospholipid N-methyltransferases and named *pmtA* and *pmtX_2* (Minder *et al.*, 2001, Hacker *et al.*, 2008). A symbiotic defect of Δ*pmtA* mutant, already observed in *G. max* nodules is also visible on *A. afraspera* ones, thus extending the symbiotic phenotype on another host plant (Minder *et al.*, 2001). A low shoot/root ratio is observed compared to the plant elicited with the WT strain, displaying a nutritional stress. Nodules are much less pink and bacteroids even if alive, seem to be much less dense than in WT-elicited symbiotic cells (Figure S6). Despite its high expression in bacteroids, no symbiotic

defects could be identified on both host plants for the Δ*pmtX_2* mutant, as observed for the orthologs of this gene cluster in *Bradyrhizobium* sp. ORS285 (Hacker *et al.*, 2008; Lamouche *et al.*, 2018; Figure S6).

Taken together, these data show that, regardless of their host, 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 expresses its symbiotic program within this peculiar host as it would do in soybean, its original host. Thus, metabolic dysfunction of these nodules does not seem to come from a bacterial defect to express the symbiotic program.

Host-specific adaptation

Comparison of the *A. afraspera* and *G. max* rescued bacteria revealed significant differences in these proteomes and transcriptomes. At the transcriptomic level, 935 DEG could be identified between 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 DEG in *A. afraspera* nodules and the presence of 5 newly annotated CDS amongst the 20 most induced DEG in *G. max* nodules (Supplementary Table 1).

At the proteomic level, 414 proteins were differentially accumulated between the two bacteroid types (171 *A. afraspera*>*G. max* and 243 *G. max*>*A. afraspera*). This is the case of proteins directly involved in microoxic respiration, such as FixA, FixB and FixC, and in nitrogen fixation, such as NifH and NifK, which were found to be more abundant in *A. afraspera* than in *G. max* (Table 2) despite the apparent lack of efficiency in *A. afraspera*. Additionally, three GroEL chaperonins (GroEL1, 5 and 6) were over-accumulated in *A. afraspera* bacteroids and GroEL2 was specifically detected in this latter condition. Interestingly, the phenylacetic acid degradation pathway (PaaABCDEIK) was highly expressed in *A. afraspera* nodules, as well as a yet uncharacterized cluster of genes putatively involved in toluene degradation (blr3675, 3678-3680). Also, several proteins related to stress like proteases were highly accumulated in *A. afraspera* nodules (Table 2).

Label	gene name	EugenePP annotation	mean SC AA	GM	YM	gene mutation
Chaperonins						
bll2059	*groEL_1*	IPR002423:Chaperonin Cpn60/TCP-1 family	279,3^{a}	250,8^{b}	80,5^{c}	
bll2060	*groES_1*	IPR020818:Chaperonin Cpn10	17,3^{a}	14,3^{b}	1,8^{c}	
blr3683	*groEL_2*	IPR002423:Chaperonin Cpn60/TCP-1 family	12^{a}	0^{b}	0^{b}	
blr5626	*groEL_5*	IPR002423:Chaperonin Cpn60/TCP-1 family	188,3^{a}	153,3^{b}	145,5^{c}	
blr6979	*groEL_6*	IPR002423:Chaperonin Cpn60/TCP-1 family	72,7^{a}	65,3^{b}	55,5^{c}	
Proteases and stress response						
bll0805	bll0805	IPR012338:Beta-lactamase/transpeptidase-like	9^{b}	5,3^{c}	14,3^{a}	
bll1431	bll1431	IPR011356:Leucine aminopeptidase/peptidase B	9,7^{a}	4,3^{b}	8,8^{a}	
blr2591	*dop*	IPR001940:Peptidase S1C	15,7^{a}	0,3^{b}	0^{b}	
blr2860	blr2860	IPR006143:RND efflux pump, membrane fusion protein	1,7^{b}	1^{c}	5,3^{a}	
blr3130	blr3130	IPR001940:Peptidase S1C	31^{a}	15^{b}	9^{c}	
bll3903	bll3903	IPR006143:RND efflux pump, membrane fusion protein	9,7^{a}	5,5^{b}	0,8^{c}	
blr5449	blr5449	IPR002933:Peptidase M20	5,3^{a}	3,8^{b}	0,8^{c}	
blr6174	*lon_2*	IPR014252:Sporulation protease LonC	12,3^{a}	7^{b}	12,8^{a}	
bll6433	bll6433	IPR012899:LTXXQ motif family protein	20,7^{a}	9,5^{b}	0^{c}	
blr7274	blr7274	IPR001940:Peptidase S1C	9,7^{a}	2,8^{b}	0^{c}	
Phenylacetic acid degradation						
blr2891	*paaA*	IPR007814:Phenylacetic acid catabolic	6^{a}	0,3^{c}	3,8^{b}	
bsr2892	*paaB*	IPR009359:Phenylacetic acid degradation B	4^{a}	0^{c}	3,5^{b}	
blr2893	*paaC*	IPR007814:Phenylacetic acid catabolic	8^{a}	1,5^{c}	4,5^{b}	
blr2894	*paaD*	IPR011883:Phenylacetate-CoA oxygenase, PaaJ subunit	2^{a}	0^{b}	0,8^{ab}	
blr2895	*paaE*	IPR001221:Phenol hydroxylase reductase	6^{a}	0,3^{b}	5^{a}	
blr2897	*paaK*	IPR011880:Phenylacetate-CoA ligase	2,7^{a}	0,3^{b}	4,3^{a}	
Uncharacterized alcohol degradation cluster						
blr3675	blr3675	IPR002085:Alcohol dehydrogenase superfamily, zinc-type	24,3^{a}	0^{b}	0^{b}	Yes
blr3678	blr3678	IPR000951:Phthalate dioxygenase reductase	5^{a}	0^{b}	0^{b}	
blr3679	blr3679	IPR003430:Methane/phenol/toluene hydroxylase	4^{a}	0^{b}	0^{b}	
blr3680	blr3680	IPR003454:Monooxygenase component MmoB/DmpM	4^{a}	0^{b}	0^{b}	
Nitrogen fixation and symbiotic islands						
blr1721	hupL_1	IPR029014:NiFe hydrogenase-like	4,3^{a}	3^{b}	0^{c}	
blr1744	*nifK*	IPR005976:Nitrogenase molybdenum-iron protein beta chain	56,7^{a}	42,8^{b}	3,3^{c}	
blr1747	*nifX*	IPR013480:Nitrogen fixation protein NifX	7,3^{a}	6^{b}	0^{c}	
bll1754	bll1754	hypothetical protein	13^{a}	11,8^{b}	0^{c}	
blr1756	*nifS_1*	IPR006234:O-succinylhomoserine sulfhydrylase	19^{a}	14,5^{b}	0^{c}	
blr1769	*nifH*	IPR000392:Nitrogenase iron protein, subunit NifH/Protochlorophyllide reductase, subunit ChlL	87,3^{a}	78^{b}	32^{c}	
blr1773	*fixB*	IPR001308:Electron transfer flavoprotein, alpha subunit	20,7^{a}	17,3^{b}	0^{c}	
blr1774	*fixC*	IPR002938:Monooxygenase, FAD-binding	37^{a}	27,8^{b}	1^{c}	
bsr1775	*fixX*	IPR012206:Ferredoxin-like, FixX	5,7^{a}	4,3^{b}	0^{c}	
bll1777	*ahpC_1*	IPR012336:Thioredoxin-like fold	12,7^{a}	9^{b}	0^{c}	
bll1791	bll1791	IPR027417:P-loop containing nucleoside triphosphate hydrolase	29,3^{a}	17,8^{b}	5^{c}	
blr1852	blr1852	IPR003721:Pantoate-beta-alanine ligase	14,3^{a}	6^{b}	0^{c}	
blr1880	blr1880	IPR019941:Transcription regulator LuxR, chaperone HchA-associated	7^{a}	1^{b}	0^{c}	Yes
blr1964	blr1964	IPR013319:Glycoside hydrolase, family 11/12	4^{a}	2^{b}	0^{c}	
blr2036	*fixR_1*	IPR002347:Glucose/ribitol dehydrogenase	19^{a}	12,8^{b}	0^{c}	
blr2037	*nifA*	IPR010113:Nif-specific regulatory protein	10^{a}	6^{b}	0^{c}	
blr2038	*fixA*	IPR012255:Electron transfer flavoprotein, beta subunit	24,7^{a}	21,3^{b}	0^{c}	
bll2063	*nrgC*	IPR009075:Acyl-CoA dehydrogenase/oxidase C-terminal	30^{a}	17^{b}	0^{c}	
bll2065	*icfA_1*	IPR001765:Carbonic anhydrase	7,7^{a}	4^{b}	0^{c}	
blr2108	blr2108	IPR000873:AMP-dependent synthetase/ligase	14,7^{a}	2,5^{b}	0^{b}	yes
blr2131	blr2131	IPR025700:L-lysine 6-monooxygenase/L-ornithine 5-monooxygenase	14,7^{a}	12,8^{b}	0^{c}	
blr2136	blr2136	IPR006234:O-succinylhomoserine sulfhydrylase	8,7^{a}	6^{b}	0^{c}	
blr2144	CYP112	IPR001128:Cytochrome P450	20,7^{a}	12,5^{b}	0^{c}	
blr2145	CYP114	IPR001128:Cytochrome P450	10,3^{a}	6,3^{b}	0^{c}	
blr2146	blr2146	IPR002347:Glucose/ribitol dehydrogenase	8^{a}	4,5^{b}	0^{c}	
Uptake hydrogenase						
bll6941	*hupL_2*	IPR001501:Nickel-dependent hydrogenase, large subunit	26^{a}	16^{b}	0^{c}	
bll6942	*hupS_2*	IPR001821:[NiFe]-hydrogenase, small subunit	9,7^{a}	5,3^{b}	0^{c}	
Redox regulation						
blr0594	*trxA_1*	IPR005746:Thioredoxin	7,7^{a}	1^{b}	8,8^{a}	
bll1317	bll1317	IPR012336:Thioredoxin-like fold	8^{b}	3,3^{c}	14,8^{a}	

Table 2. Selection of *B. diazoefficiens* USDA110 proteins specific for 14dpi *A. afraspera* nodules. Mean SC are the mean counted spectra among 4 replicates. Superscript letters represent significant differences after ANOVA and post hoc Tukey tests ($p<0.05$).

Oppositely, functions involved in phosphate uptake (ie. PstS, PstB, PhoD and PhoUB) were accumulated in *G. max* nodules (Table 3). Also, two αβ hydrolases (blr6576 and bll6577) and a TonB dependent receptor like protein (blr2460) were over-accumulated in a *G. max*-specific

manner. Finally, cell cycle functions such as the replication initiator DnaA and the chromosome maintenance protein Smc were present at higher level in *G. max* bacteroids as compared to *A. afraspera* bacteroids, suggesting differential influence of the two host on the bacteroid cell cycle (Table 3).

Functional analysis was undertaken for 7 functions that were differentially regulated between hosts and none of the corresponding mutants displayed any symbiotic phenotype (Tables 2-3). However, bll1944 protein highly accumulated in soybean in WT seems to have an intermediate phenotype. Indeed, the Δ1944 mutant displays lower efficiency and nodule number in *G. max* compared to the WT, whereas no such differences are observed on *A. afraspera* host (Figure S5). At the transcriptional level, genes involved in Calvin-Benson-Bassham pathway were induced in *A. afraspera* bacteroids. However, a polar mutant Ω*cbbP* of these genes and the mutant Δ*cbbR* impaired for the positive regulator of these genes display no phenotypes on both host plants (Figure S7; Masuda *et al.*, 2017).

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

It has been recently described that, during the interaction between *A. afraspera* and *B. diazoefficiens* USDA110, TBD is occurring to a very limited extent, if any, a situation reminiscent to the behavior of this strain in *G. max* where no TBD occurs (Barrière *et al.*, 2017). However, the -omics data show differential accumulation of DnaA protein in the two hosts which might reflect subtle differences in the affectation of the bacterial cell cycle between the two plants.

To better characterize the physiological state of USDA110 in *A. afraspera*, we analyzed bacteroid differentiation features by estimating cell size, bacteroid DNA content and cell viability of 14 dpi bacteroids extracted from *G. max* and *A. afraspera* nodules. As a reference, the strain ORS285 which bacteroids differ drastically in cell size and DNA content as compared to their free-living counterpart was used. The interaction between *A. afraspera* and *Bradyrhizobium* sp. ORS285 leads to the terminal differentiation of bacteroids into elongated polyploid cells (Bonaldi *et al.*, 2011; Guefrachi *et al.*, 2015).

Label	gene name	EugenePP annotation	mean SC AA	GM	YM	Mutation
cell cycle functions						
bll0830	*dnaA*	IPR020591:Chromosomal replication control, initiator DnaA-like	$0,7^{b}$	$4,5^{a}$	2^{b}	
bll7146	*ftsH*	IPR003959:ATPase, AAA-type, core	$0,3^{b}$	$5,8^{a}$	9^{a}	
Iron aquisition						
bll0797	*fur*	IPR002481:Ferric-uptake regulator	0^{b}	$2,3^{a}$	$1,3^{a}$	
bll2460	bll2460	IPR010105:TonB-dependent siderophore receptor	0^{b}	$26,8^{a}$	0^{b}	Yes
Phosphate /phosphonate metabolism						
blr1091	*pstS*	IPR005673:ABC transporter, phosphate-binding protein PstS	19^{b}	$25,8^{a}$	$5,5^{c}$	
blr1094	*pstB*	IPR005670:Phosphate transport system permease protein 1	$2,7^{b}$	8^{a}	0^{c}	
blr1095	*phoU*	IPR026022:PhoU domain	$4,3^{b}$	6^{a}	$0,3^{c}$	
blr1096	*phoB*	IPR011006:CheY-like superfamily	$5,7^{b}$	$7,8^{a}$	$1,3^{c}$	
blr1227	*phnM_1*	IPR012696:Phosphonate metabolism PhnM	0^{b}	2^{a}	0^{b}	
bll7946	*phoD*	IPR005770:Phosphate/phosphite/phosphonate ABC transporter, periplasmic protein	$2,3^{b}$	$8,8^{a}$	0^{c}	
blr7948	blr7948	IPR017694:Phosphonate metabolim protein, transferase hexapeptide repeat family	0^{b}	2^{a}	0^{b}	
blr7949	*phnM_2*	IPR012696:Phosphonate metabolism PhnM	$0,7^{b}$	$2,8^{a}$	0^{b}	
Branched aminoacid metabolism						
bll0144	bll0144	IPR000709:Leu/Ile/Val-binding protein	0^{b}	3^{a}	$1,5^{ab}$	
bll0887	bll0887	IPR000709:Leu/Ile/Val-binding protein	$11,3^{c}$	25^{a}	$17,8^{b}$	
bll0912	bll0912	IPR000709:Leu/Ile/Val-binding protein	$0,7^{c}$	$2,3^{b}$	$10,8^{a}$	
blr0998	blr0998	IPR000709:Leu/Ile/Val-binding protein	$1,7^{b}$	6^{a}	6^{a}	
blr1052	blr1052	IPR000709:Leu/Ile/Val-binding protein	$0,3^{c}$	$3,5^{b}$	$10,8^{a}$	
blr1147	blr1147	IPR000709:Leu/Ile/Val-binding protein	0^{c}	$1,3^{b}$	$5,3^{a}$	
blr1448	blr1448	IPR000709:Leu/Ile/Val-binding protein	$2,7^{b}$	$10,5^{a}$	0^{c}	
blr3353	blr3353	IPR000709:Leu/Ile/Val-binding protein	$0,7^{b}$	$7,8^{a}$	$0,3^{b}$	
blr3825	blr3825	IPR000709:Leu/Ile/Val-binding protein	$1,3^{c}$	6^{b}	10^{a}	
blr4039	blr4039	IPR000709:Leu/Ile/Val-binding protein	$7,7^{c}$	14^{b}	$16,8^{a}$	
bll4057	bll4057	IPR000709:Leu/Ile/Val-binding protein	7^{c}	$7,8^{b}$	$27,3^{a}$	
blr5675	blr5675	IPR000709:Leu/Ile/Val-binding protein	$8,7^{c}$	16^{b}	$22,8^{a}$	
bll5953	bll5953	IPR000709:Leu/Ile/Val-binding protein	$5,7^{c}$	$10,3^{b}$	15^{a}	
blr5966	blr5966	IPR000709:Leu/Ile/Val-binding protein	0^{c}	$1,3^{b}$	$9,8^{a}$	
blr6078	blr6078	IPR000709:Leu/Ile/Val-binding protein	0^{b}	$3,3^{a}$	$3,8^{a}$	
bll6899	bll6899	IPR000709:Leu/Ile/Val-binding protein	4^{b}	7^{a}	$0,8^{c}$	
blr7827	blr7827	IPR000709:Leu/Ile/Val-binding protein	$7,3^{c}$	$13,3^{b}$	$19,8^{a}$	
blr7828	blr7828	IPR000709:Leu/Ile/Val-binding protein	$0,7^{c}$	$5,3^{b}$	$7,3^{a}$	
Enveloppe modification						
bll6634	pmtX_2	IPR029063:S-adenosyl-L-methionine-dependent methyltransferase	0^{b}	$4,8^{a}$	0^{b}	
bll2362	bll2362	IPR003856:Lipopolysaccharide biosynthesis	0^{c}	$9,5^{a}$	3^{b}	
blr7131	blr7131	IPR002477:Peptidoglycan binding-like	$0,3^{c}$	$4,8^{b}$	$8,3^{a}$	
blr4219	blr4219	IPR005490:L,D-transpeptidase catalytic domain	0^{b}	$3,3^{a}$	$3,8^{a}$	
Nitrogen fixation and symbiotic islands						
blr1745	*nifE*	IPR005973:Nitrogenase MoFe cofactor biosynthesis protein NifE	$14,7^{b}$	$17,3^{a}$	0^{c}	
blr1770	*nifQ*	IPR006975:NifQ	$9,3^{b}$	$10,3^{a}$	0^{c}	
bll2067	*nfeC*	nodulate formation efficiency C protein	13^{b}	$14,8^{a}$	0^{c}	
bll1634	bll1634	hypothetical protein	3^{b}	$5,5^{a}$	0^{c}	
blr1686	blr1686	IPR005814:Aminotransferase class-III	$25,7^{b}$	28^{a}	$1,8^{c}$	Yes
blr1765	*fer2*	IPR012336:Thioredoxin-like fold	2^{b}	3^{a}	0^{c}	
bll1766	bll1766	IPR005618:Outer membrane protein, OmpW	2^{b}	4^{a}	0^{c}	
blr1830	blr1830	hypothetical protein	17^{b}	$27,5^{a}$	$0,3^{c}$	
blr1851	blr1851	IPR001646:Pentapeptide repeat	$3,7^{b}$	$4,8^{a}$	0^{c}	
bll1872	bll1872	IPR011250:Outer membrane protein/outer membrane enzyme PagP , beta-barrel	7^{b}	$10,5^{a}$	0^{c}	Yes
bll1875	bll1875	IPR001214:SET domain	$3,7^{b}$	$4,5^{a}$	0^{c}	
bll1944	bll1944	IPR011250:Outer membrane protein/outer membrane enzyme PagP , beta-barrel	10^{b}	$15,8^{a}$	0^{c}	Yes
blr2006	blr2006	hypothetical protein	$2,3^{b}$	$6,3^{a}$	0^{c}	
bll2007	*hemN1*	IPR004558:Coproporphyrinogen III oxidase, oxygen-independent, HemN	$4,7^{b}$	$11,3^{a}$	0^{c}	
blr2042	blr2042	IPR005123:Oxoglutarate/iron-dependent dioxygenase	$5,7^{b}$	$7,3^{a}$	0^{c}	
bll2049	*trpD_1*	IPR005940:Anthranilate phosphoribosyl transferase	$8,3^{b}$	$11,5^{a}$	0^{c}	
blr2068	blr2068	IPR006276:Cobalamin-independent methionine synthase	13^{b}	$17,8^{a}$	0^{c}	

Table 3. Selection of *B. diazoefficiens* USDA110 proteins specific for 14dpi *Glycine max* nodules. Mean SC are the mean counted spectra among 4 replicates. Superscript letters represent significant differences after ANOVA and post hoc Tukey tests ($p<0.05$).

TBD is hallmarked by the polyploidy state of bacteroids. However, USDA110 shows only two peaks at 1C and 2C in *G. max*, which is similar to cycling cells in the bacterial culture

sample (Figure 3A-B). Similar results were obtained for most of the experiments performed on *A. afraspera*, however, a third peak at 3C was occasionally observed in *A. afraspera* bacteroids and to a lower extent, a peak at 4C. Nonetheless, ORS285 reaches higher levels of endoreduplication with peaks at 6C and more (average of 8C according to Czernic *et al.*, 2015). Thus, regarding the ploidy level USDA110 displays limited features of a typical TBD in *A. afraspera* nodules.

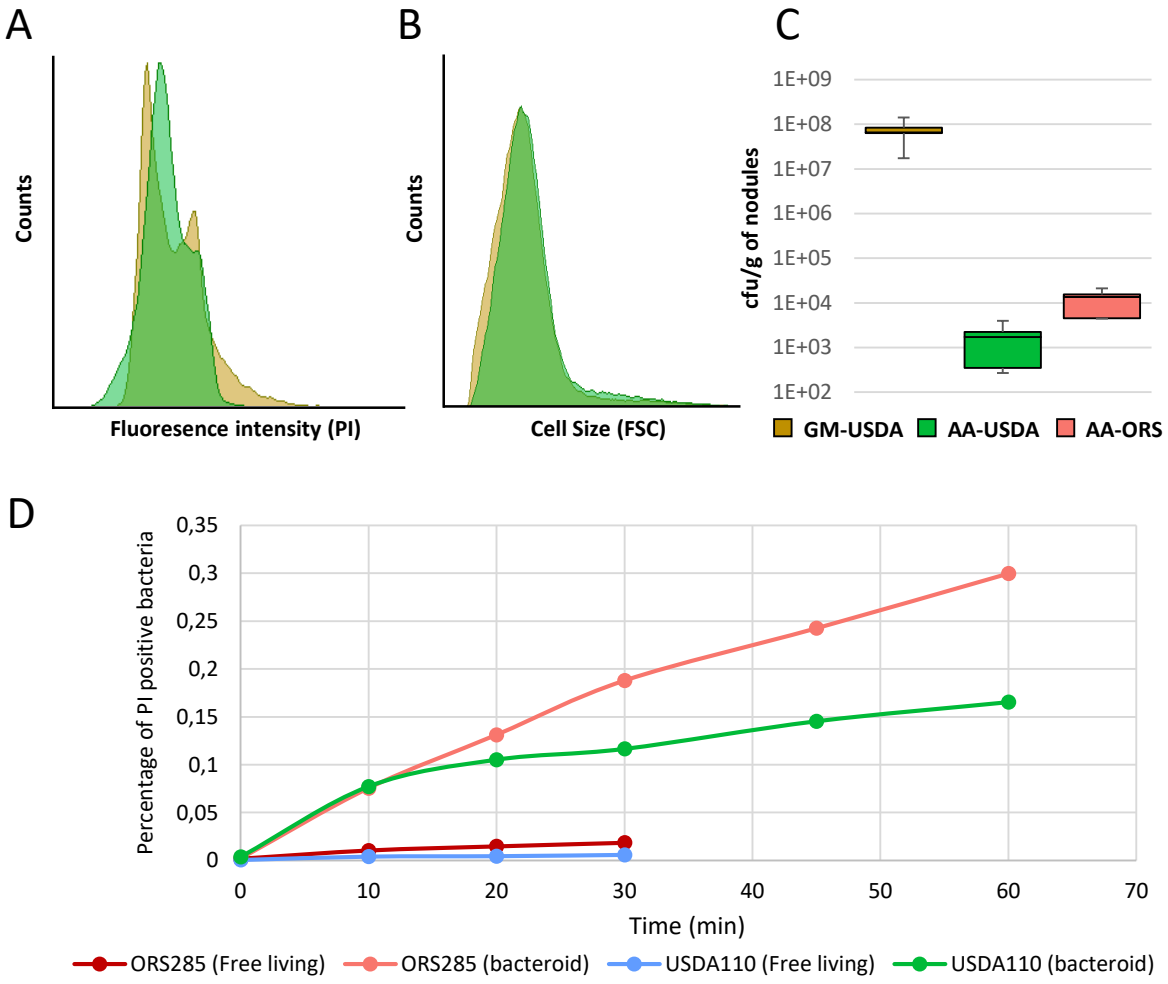


Figure 3. *B. diazoefficicens* USDA110 displays atypical bacteroid differentiation features in *A. afraspera* nodules. A, B. Assessment of the ploidy level and size of soybean and *A. afraspera* extracted USDA110 bacteroids with flow cytometry. C. Viability of soybean and *A. afraspera* extracted bacteroids at 14 dpi determined by comparing the number of bacteroid per mg nodule counted by flow cytometry and plating followed by cfu counting after 5 days of growth. D. Permeability of *A. afraspera* extracted USDA110 or ORS285 bacteroids at 14 dpi compared to free-living bacteria (preliminary results).

To analyze the viability of the bacteroids, nodule extracted bacteroids were plated and the colony forming units (cfu) were counted (Figure 3C). In *G. max* where TBD is not occurring,

USDA110 formed 10^8 colonies/mg nodule (~100% survival). Oppositely, ORS285 formed only 10^4 colonies/mg nodule in *A. afraspera* where TBD occurs (~1% survival). Interestingly, USDA110 formed 2.10^3 colonies/mg nodule in *A. afraspera* (~0,2% survival), indicating that, despite almost no variation in cell morphology and a moderate increase in ploidy level as compared to their free-living relatives, USDA110 bacteroids undergo a *bona fide* terminal differentiation in *A. afraspera*.

Loss of membrane integrity is a hallmark of TBD that likely strongly contributes to the loss of viability of bacteroids. Preliminary results of time course analysis of propidium iodide (PI) uptake by nodule extracted bacteroids and the corresponding culture controls were performed and showed that the envelope of USDA110 bacteroids is permeable in *A. afraspera*, like the one of ORS285 (Fig 3D). In contrast to bacteroids, the cells collected from cultures never took up the PI, indicating that their membranes were not destabilized. The fluorescence level of both USDA110 and ORS285 bacteroids increased over time but with a delay in USDA110, suggesting that the envelope of USDA110 bacteroids might be less permeable than the one of ORS285.

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 (Barrière *et al.*, 2017). Here we show that USDA110 reduced endoreduplication level and limited cell elongation occur in terminally differentiated bacteroid that fix nitrogen in a suboptimal way. Accordingly, the protein level of DnaA, the genome replication initiator, was higher at 14dpi in soybean than in *A. afraspera* confirming that polyploidization did not occur in this latter host. In soybean, bacteroids may still divide within symbiosomes, whereas in *A. afraspera*, they already reached the final stage of differentiation as with ORS285 nodulating this host at the same timepoint. Such unusual terminal bacteroid differentiation is reminiscent of the *Glycyrrhiza uralensis* situation. This plant of the Inverted Repeat Lacking Clade was reported to have and express NCR peptides (Montiel *et al.*, 2017). But 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 (Crespo-Rivas *et al.*, 2016). 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 (Price *et al.*, 2015). In these two plants, there is a form of strain specificity but no uncoupling of the terminal aspect and of the morphological change and endoreduplication as found in USDA110-*A. afraspera*.

The surprising differentiation of USDA110 in *A. afraspera* nodules rise questions about the molecular mechanisms supporting this phenomenon. We consider two possible hypotheses to explain this observation: USDA110 might be more sensitive to differentiation agents than ORS285 does and thus be too 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 elongation and polyploidization features.

Accordingly, the application of NCR peptides has very limited effect on USDA110 as compared to *S. meliloti* and to other plant-associated bacteria (Tiricz *et al.*, 2013; Barrière *et al.*, 2017). NCR insensitivity may be due to the thick hopanoid layer that is present in USDA110 outer membrane, as the hopanoid biosynthesis mutant *hpnH* shows symbiotic defects in *A. afraspera* but not in *G. max* (Kulkarni *et al.*, 2015).

Terminal differentiation is associated with specific stress response

The terminal aspect of bacteroid differentiation is not observed in *G. max*, the original host of USDA110, and is associated with a higher accumulation of stress markers in the bacteria extracted from *A. afraspera* nodules as compared to *G. max* ones. This is the case of 3 proteases (Dop, lon_2 and blr3130) and one chaperonin (GroEL_2). Interestingly, genes encoding these stress related proteins are not part of the general stress response (GSR) and the PhyR/EcfG regulon is not activated *in planta* and does not show differential expression between hosts (Gourion *et al.*, 2009). 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* (Lamouche *et al.*, 2018), suggesting strain specificity in the activation of the GSR in *A. afraspera*. Interestingly, similar induction of proteases and chaperonins have been reported in NCR-treated *S. meliloti* cultures (Tiricz *et al.*,

2013), indicating that this response may be rather linked to the perception of *A. afraspera* NCR-like peptides in USDA110.

Correlation between bacteroid differentiation features and symbiotic efficiency for the plant.

Terminal bacteroid differentiation is associated to the massive production of symbiotic antimicrobial peptides such as NCR, NCR-like and CAPE peptides in different plants (Mergaert *et al.*, 2003, Czernic *et al.*, 2015; Karmakar *et al.*, 2018). They represent ~10% of nodule transcriptomes in *Medicago truncatula* and thereby a huge energetic cost for the plant rising questions about the benefits of such strategy (Guefrachi *et al.*, 2014). The reason why TBD appeared independently in different legume clades remains unclear but suggests 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 (Sen and Weaver, 1981; Oono and Denison, 2010; Lamouche *et al.*, 2018. Interestingly, the data reported here are in agreement with this hypothesis as the more differentiation features the bacteroids display, as noticed with ORS285, a photosynthetic strain originally isolated from *A. afraspera* nodules (Molouba *et al.*, 1999), the more important improvement of plant growth is observed. However, the simultaneous analysis of bacteroid differentiation level and symbiotic performance of different *Aeschynomene-Bradyrhizobium* interactions has shown that there is no clear link between the extent of bacteroid differentiation and the symbiotic efficiency of the plant-bacterium couple (Lamouche *et al.*, under revision in Frontiers in Plant Science).

Underestimation of terminal differentiation without morphological changes?

Interestingly, in *A. afraspera* nodules, not all ORS285 bacteria are morphologically differentiated and bacteria displaying contrasted morphologies are not colonizing the same zones. A central nodule zone harbors differentiated bacteroids expressing nitrogen fixation genes whereas a nodule cap contains undifferentiated rhizobia that do not express those genes (Renier *et al.*, 2011; Bonaldi *et al.*, 2011). Despite USDA110-triggered nodules do not contain this second zone (ie. the cap), a similar recovery of viable cells was obtained for both ORS285 and USDA110 in *A. afraspera*. This similar level suggests that *A. afraspera* hosted ORS285 bacteria that do not fix nitrogen and that harbor undifferentiated morphology could be

terminally differentiated. In the IRLC clade, the extent of bacteroid differentiation was correlated to the size of the cationic NCR peptides repertoire (Montiel *et al.*, 2017).

Nevertheless, the discovery of bacteroids that are terminally differentiated without any obvious morphological changes opens the possibility that the terminal nature of bacteroid differentiation might have been underestimated in other legume plants, in which terminal differentiation might occur without any morphological changes.

Interest to combine transcriptomics and proteomics

Here we reported the combined use of transcriptomics and of proteomics to investigate the physiology of bacteroids in *G. max* and in *A. afraspera* host. Those two approaches are complementary. Indeed, while detection of a protein is more relevant with respect to the physiological state than the abundance of the corresponding transcript, the detection limits of the current commonly employed proteomic technics only allow partial characterization of the proteomes. In contrast, today's RNA-seq approaches reached the depth of sequencing to obtain exhaustive expression data for all defined genes and other genomic objects (such as ncRNA or tRNA) and even in some cases to identify new ones as we did in this study.

USDA110 transcriptomics data in the perspective of previously described regulons

USDA110 is one of the best characterized rhizobial strain with transcriptomics responses to various stimuli described as well as many regulons (diCenzo *et al.*, 2018). The study of gene regulation, as coexpression of full regulons/stimulons can be easily searched in the transcriptomes reported here. We analyzed the behavior of these gene networks in our dataset (Supplementary Table 2). 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 14dpi, when the nodule is formed and functioning, the genistein stimulon, which comprises the NodD1, NodVW, TtsI and LafR regulons, is no more activated *in planta*. Nevertheless, 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*. 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 vanillate degradation pathway was observed in the transcriptome of *Bradyrhizobium* sp. ORS285 in *A. afraspera* and *A. indica* nodules (Lamouche *et al.*, 2018), 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 (Teufel *et al.*, 2010). Similarly, enzymes of another pathway involved in phenolic compound degradation (blr3675-blr3680) are accumulated in *A. afraspera* nodules.

Microoxic Sanction?

The accumulation of symbiotic proteins in the *A. afraspera* host, such as proteins of the nitrogenase complex, as compared to *G. max* might seem contradictory with the reduced symbiotic capacity of USDA110 in *A. afraspera*. In *B. diazoefficiens*, two regulatory cascades, RegSR-NifA and FixLJ-$FixK_2$/K_1, control the expression of nitrogen fixation genes and genes required for microoxic life, respectively (Terpolilli *et al.*, 2012). Although the regulons are not differentially expressed between hosts according to our criteria in the transcriptomic dataset, the proteomics data show interesting features of these regulons. We quantified only two FixJ regulated functions (blr3678 and blr3683-GroEL4), and they indeed show higher accumulation in *A. afraspera* as compared to *G. max* nodules. Accumulation observed in *A. afraspera* bacteria might reflect the concentration of free oxygen in the nodule of this species which would not be optimal for USDA110. However, when comparing the sum of reads for the nitrogenase complex (NifHDK), we observe that in *A. afraspera* nodules, we saw that they represent ~12% of the dataset for ORS and only ~6% for USDA110 (Lamouche *et al.*, 2018). Nevertheless, the concentration of free oxygen in the nodules is strongly dependent on the abundance of a plant protein, the leghemoglobin.

In line with this, ascorbate has been shown to reduce leghemoglobin and its addition to plant medium can induce a higher nitrogenase activity in the nodules (Ross *et al.*, 1999). Thus, as we observe an increase in ascorbate in *A. afraspera*/USDA110 nodules, together with an increase in the accumulation of nitrogenase proteins, it is likely that the host plant is diminishing the

oxygen pressure in the nodule. In previous studies, the decrease in nodule oxygen level has been suggested to be the mean of sanctioning non-cooperative bacteria (Kiers *et al.*, 2003). In the case where TBD does not reach its final stage, as shown here in the *A. afraspera*/USDA110 interaction, the plant may be sanctioning the suboptimal nitrogen fixer. An alternative explanation would be that the possible decrease in nodule oxygen pressure triggered by the host would be a way to improve the nitrogen fixation rate of a suboptimal interaction. Indeed, if nitrogen fixation is not reaching the expected rate, the plant may decrease oxygen pressure to restore the best conditions for the nitrogenase to function.

We did not properly assess the scenario of host sanction, however, by plating *A. afraspera* nodule extracted bacteria, we could not identify any difference between the survival rate of ORS285 and the one of USDA110, suggesting that the plant did not affect bacterial survival through this possible modulation of the nodule oxygen level. Moreover, this mechanism of host sanction and the means that are used to sanction do not seem to be universal in rhizobium legume symbiosis (Gubry-Rangin *et al.*, 2010).

Conclusion

USDA110 is a major model in the legume-rhizobium symbiosis, mainly thanks to its interaction with *G. max*, the most produced legume worldwide. Although transcriptomic and proteomic studies have been conducted in this strain in symbiosis with various hosts, 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 indeed functional even if nitrogen fixation and plant benefits remain sub-optimal, as hallmarked by the nodule metabolomic profiles.

Nevertheless, terminal bacteroid differentiation is taking place as bacterial viability is impaired in USDA110 bacteroids, whereas morphological changes are not observed and cell cycle switch to endoreduplication seem to be very limited. We also show that the bacterial symbiotic program is taking place 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 its unusual host *A. afraspera*, 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*) will be a challenging step due to the differences in the gene repertoire of those bacteria, but it could allow a broader comparison of bacterial physiology in cases where TBD results in different bacterial outputs.

Acknowledgments

F.L. and Q.N were supported by a PhD fellowship from the Université Paris-Sud. We would like to thank Hans-Martin Fischer and Evelyne Bauer from ETH Zurich for Ω*cbbP* and Δ*cbbR* mutant strains. 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, grant n° ANR-13-BSV7-0013 and used resources from the National Office for Research, Development and Innovation of Hungary, grant n° 120120 to A.K. 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.

Supplemental material

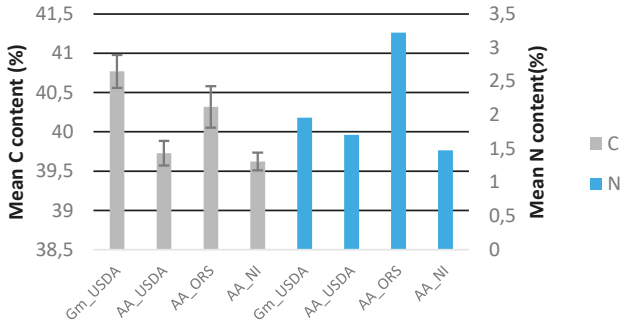


Figure S1. Nitrogen and carbon content in aerial parts of the plants were determined by EA-IRMS. Gm: *Glycine max*, AA: *Aeschynomene afraspera*, ORS: inoculated by *Bradyrhizobium* sp. ORS285, USDA: inoculated by *B. diazoefficiens* USDA110, NI: Not inoculated

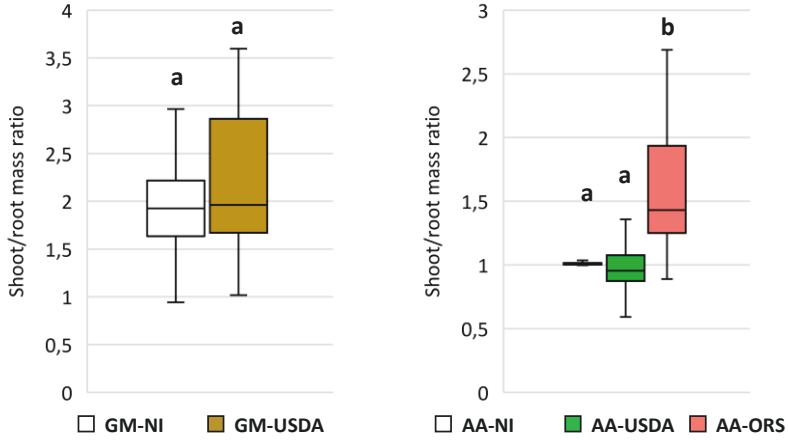


Figure S2. Nutritional status of 14 dpi plants determined by the shoot/root mass ratios. Gm: *Glycine max*, AA: *Aeschynomene afraspera* ,ORS: inoculated by *Bradyrhizobium* sp. ORS285, USDA: inoculated by *B. diazoefficiens* USDA110, NI: Not inoculated. Letters represent significant differences after t-test or ANOVA and post hoc Tukey tests ($p<0.05$).

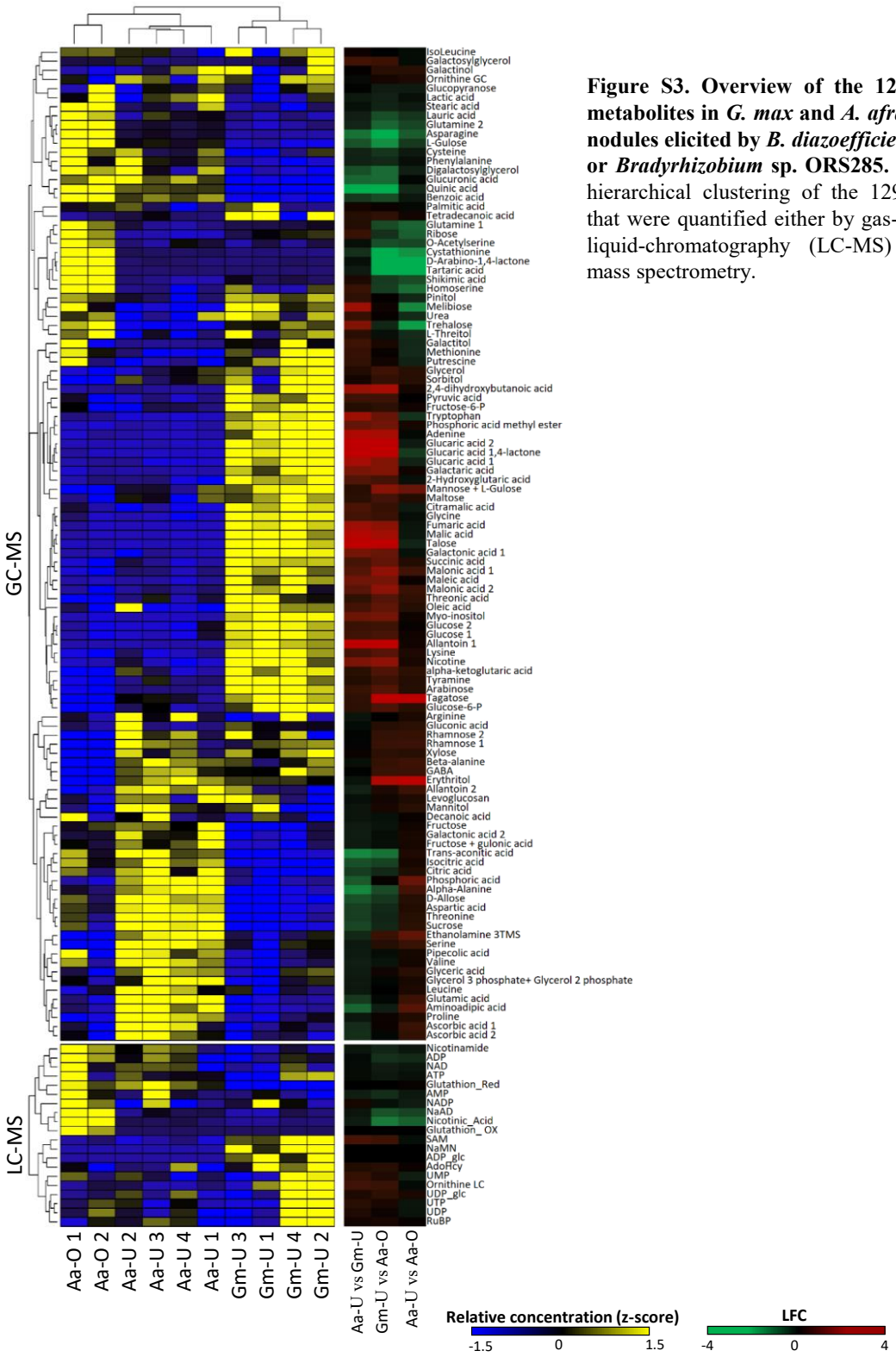


Figure S3. Overview of the 129 quantified metabolites in *G. max* and *A. afraspera* whole nodules elicited by *B. diazoefficiens* USDA110 or *Bradyrhizobium* sp. ORS285. Heatmap and hierarchical clustering of the 129 metabolites that were quantified either by gas- (GC-MS) or liquid-chromatography (LC-MS) coupled to mass spectrometry.

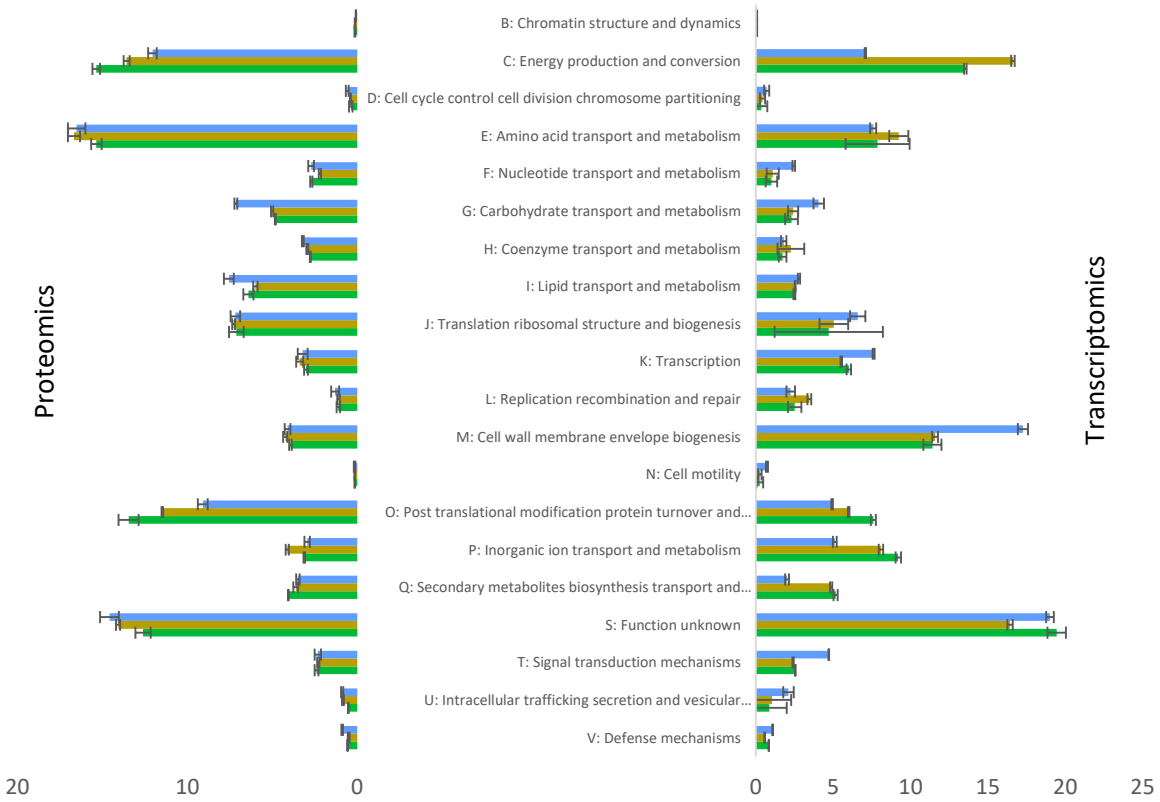


Figure S4. General overview of the datasets using COG classification. Repartitions of the assigned spectra (left panel) and reads (right panel) among COG classes in the three conditions.

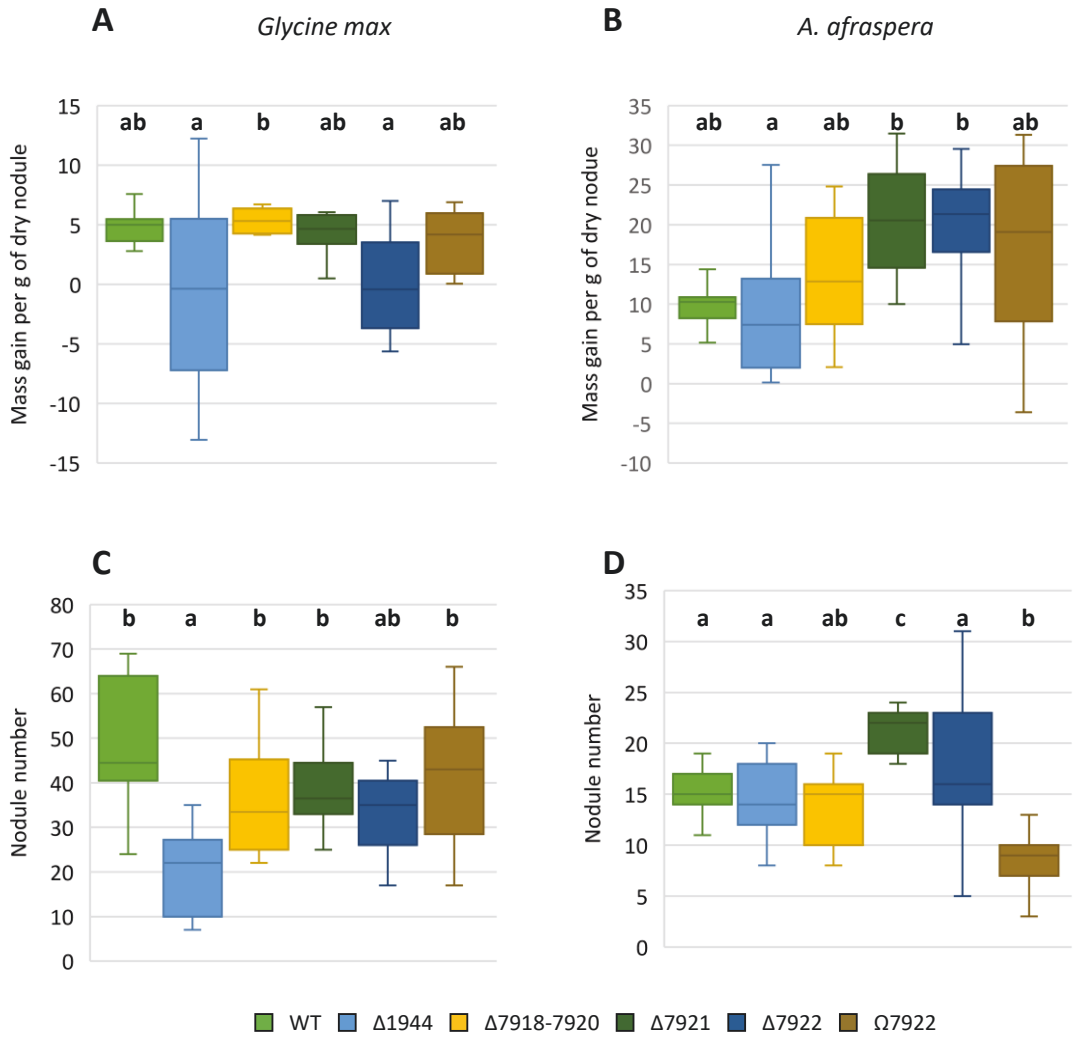


Figure S5. Functional studies of selected genes displaying no or little symbiotic phenotypes. (A, B). Comparative symbiotic mass gain of *B. diazoefficiens* USDA110 mutants compared to WT strain in *Glycine max* (A) and *A. afraspera* (B) nodules. (C,) Nodule number elicited by *B. diazoefficiens* USDA WT and mutant strains in *Glycine max* (C) and *A. afraspera* (D). Letters represent significant differences after ANOVA and post hoc Tukey tests ($p<0.05$).

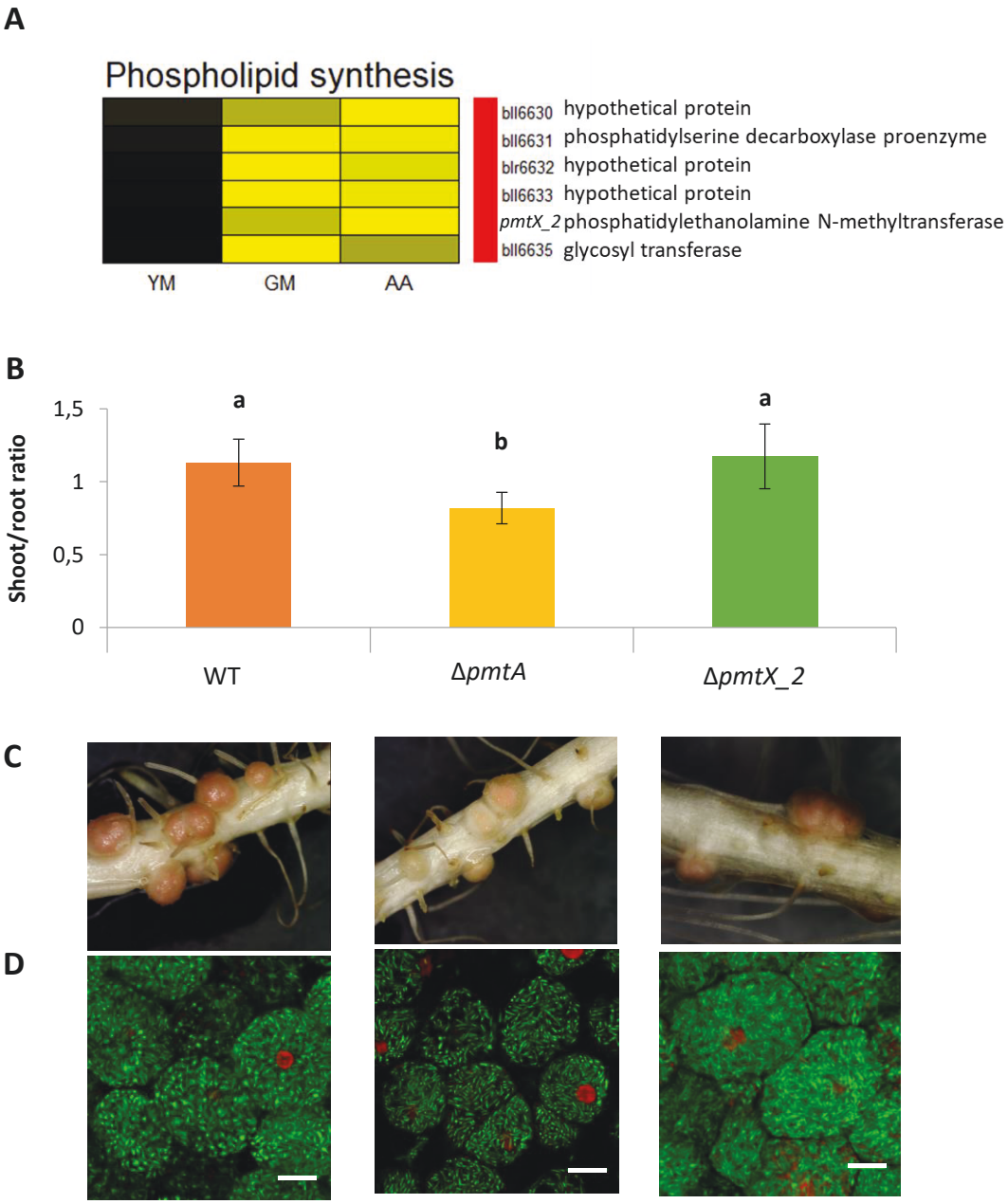


Figure S6. Functional studies of genes involved in phospholipid synthesis. A. Heatmap of a gene cluster upregulated in both bacteroid conditions. The right column indicates whether the gene is DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). B. Nutritional state of mutant Δ*pmtA* and Δ*pmtX_2* elicited nodules of *A. afraspera* measured by shoot root ratio is similar as WT. C. Macroscope images of *A. afraspera* nodules, confirming the symbiotic phenotype of Δ*pmtA* mutant previously observed on *G. max* . D. Confocal images of symbiotic cells showing a reduced bacteroid density for Δ*pmtA* mutant. Letters represent significant differences after ANOVA and post hoc Tukey tests (p<0.05).

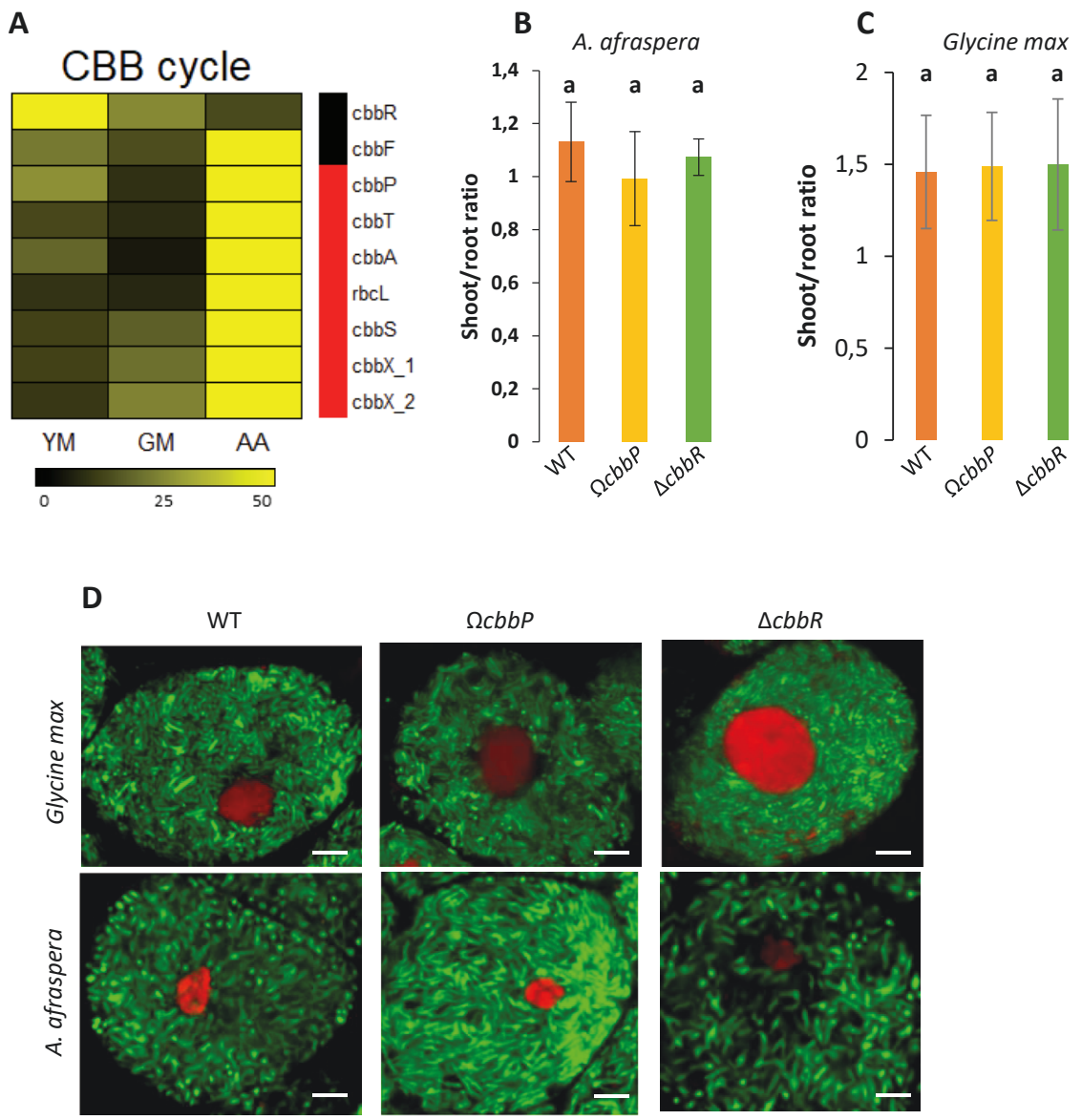


Figure S7. Functional studies of genes involved in Calvin Benson Bassham pathway. A. Heatmap of *cbb* gene expression (RNA-seq), showing their upregulation in *A. afraspera* bacteroids The right column indicates whether the gene is DEG (red) or not (black) in at least one condition (fdr<0.01 and |LFC|>1.58). B, C. Nutritional status of Ω*cbbP* and Δ*cbbR* elidited nodules of *A. afraspera* (B) and *G. max* (C) mutants measured by shoot/root ratio is similar to the WT control. D. Confocal images of symbiotic cells displaying an effective symbiosis for the tested mutants. Letters represent significant differences after ANOVA and post hoc Tukey tests ($p<0.05$).

References

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.

Alunni, B., Kevei, Z., Redondo-Nieto, M., Kondorosi, A., Mergaert, P., and 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.

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.

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.

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.

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.

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., *et al.* (2015) Convergent evolution of rndosymbiont differentiation in Dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by nodule-specific cysteine-rich peptides. *Plant Physiol* **169**: 1254–1265.

Delmotte, N., Ahrens, C.H., Knief, C., Qeli, E., Koch, M., Fischer, H.M., *et al.* (2010) An integrated proteomics and transcriptomics reference data set provides new insights into the *Bradyrhizobium japonicum* bacteroid metabolism in soybean root nodules. *Proteomics* **10**: 1391–1400.

Delmotte, N., Mondy, S., Alunni, B., Fardoux, J., Chaintreuil, C., Vorholt, J.A., *et al.* (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**: 3660–3670.

diCenzo, G.C., Zamani, M., Checcucci, A., Fondi, M., Griffitts, J.S., Finan, T.M., and Mengoni, A. (2018) Multi-disciplinary approaches for studying rhizobium – legume symbioses. *Peer J* cjm-2018-0377.

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.

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.

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.

Gourion, B., Sulser, S., Frunzke, J., Francez-Charlot, A., Stiefel, P., Pessi, G., *et al.* (2009) The PhyR-σEcfG signalling cascade is involved in stress response and symbiotic efficiency in *Bradyrhizobium japonicum*. *Mol Microbiol* **73**: 291–305.

Gubry-Rangin, C., Garcia, M., and Béna, G. (2010) Partner choice in *Medicago truncatula-Sinorhizobium* symbiosis. *Proceedings Biol Sci* **277**: 1947–51.

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**: 712.

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.

Hacker, S., Sohlenkamp, C., Aktas, M., Geiger, O., and Narberhaus, F. (2008) Multiple phospholipid N-methyltransferases with distinct substrate specificities are encoded in *Bradyrhizobium japonicum*. *J Bacteriol* **190**: 571–580.

Hernandez, G., Valdes-Lopez, O., Ramirez, M., Goffard, N., Weiller, G., Aparicio-Fabre, R., *et al.* (2009) Global changes in the transcript and metabolic profiles during symbiotic nitrogen fixation in phosphorus-stressed common bean plants. *Plant Physiol* **151**: 1221–1238.

Karmakar, K., Kundu, A., Rizvi, A.Z., Dubois, E., Severac, D., Czernic, P., *et al.* (2018) 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* MPMI-06-18-0174-R.

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.

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.

Lamouche, F., Gully, D., Chaumeret, A., Nouwen, N., Verly, C., Pierre, O., *et al.* (2018) Transcriptomic dissection of *Bradyrhizobium* sp. strain ORS285 in symbiosis with *Aeschynomene* spp. inducing different bacteroid morphotypes with contrasted symbiotic efficiency. *Environ Microbiol* **0**:.

Langella, O., Hoogland, C., Joets, J., Zivy, M., Valot, B., Jacob, D., *et al.* (2013) Management and dissemination of MS proteomic data with PROTICdb: Example of a quantitative comparison between methods of protein extraction. *Proteomics* **13**: 1457–1466.

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.

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 nnfection and colonization. *Mol Plant-Microbe Interact* **28**: 959–967.

Masuda, S., Hennecke, H., and Fischer, H.M. (2017) Requirements for efficient thiosulfate oxidation in *Bradyrhizobium diazoefficiens*. *Genes (Basel)* **8**: 1–13.

Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., and Kondorosi, E. (2003) A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiol* **132**: 161–73.

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.

Minder, A.C., De Rudder, K.E.E., Narberhaus, F., Fischer, H.M., Hennecke, H., and Geiger, O. (2001) Phosphatidylcholine levels in *Bradyrhizobium japonicum* membranes are critical for an efficient symbiosis with the soybean host plant. *Mol Microbiol* **39**: 1186–1198.

Molouba, F., Lorquin, J., Willems, A., Hoste, B., Giraud, E., Dreyfus, B., *et al.* (1999) Photosynthetic bradyrhizobia from *Aeschynomene* spp. are specific to stem-nodulated species and form a separate 16S ribosomal DNA restriction fragment length polymorphism group. *Appl Environ Microbiol* **65**: 3084–3094.

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.

Okazaki, S., Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., *et al.* (2015) Rhizobium–legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. *ISME J* **10**: 64–74.

Oono, R. and Denison, R.F. (2010) Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol* **154**: 1541–1548.

Price, P.A., Tanner, H.R., Dillon, B.A., Shabab, M., Walker, G.C., and Griffitts, 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.

Regensburger, B. and Hennecke, H. (1983) RNA polymerase from *Rhizobium japonicum*. *Arch Microbiol* **135**: 103–109.

Renier, A., Maillet, F., Fardoux, J., Poinsot, 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.

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.

Sallet, E., Gouzy, J., and Schiex, T. (2014) EuGene-PP: a next-generation automated annotation pipeline for prokaryotic genomes. *Bioinformatics* **30**: 2659–61.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., *et al.* (2010) Genome sequence of the palaeopolyploid soybean. *Nature* **463**: 178–183.

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.

Stonoha-arther, C. and 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.

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.

Terpolilli, J.J., Hood, G.A., and Poole, P.S. (2012) What determines the efficiency of N2-fixing rhizobium-legume symbioses? In, *Advances in Microbial Physiology*. Elsevier Ltd., pp. 325–389.

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**: 14390–14395.

Tiricz, H., Szücs, 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.

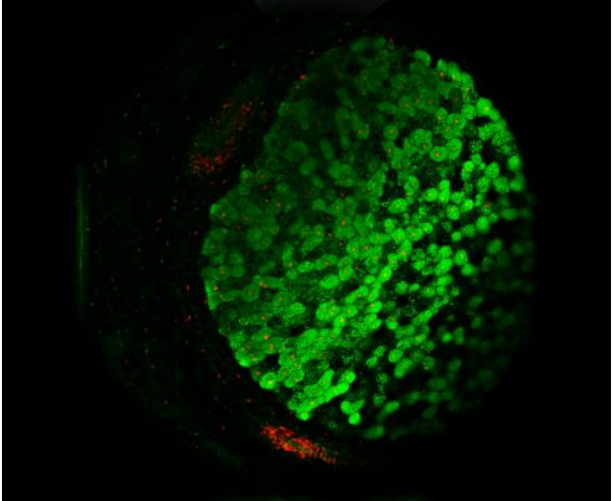
Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., *et al.* (2010) Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* **327**: 1122–1126.

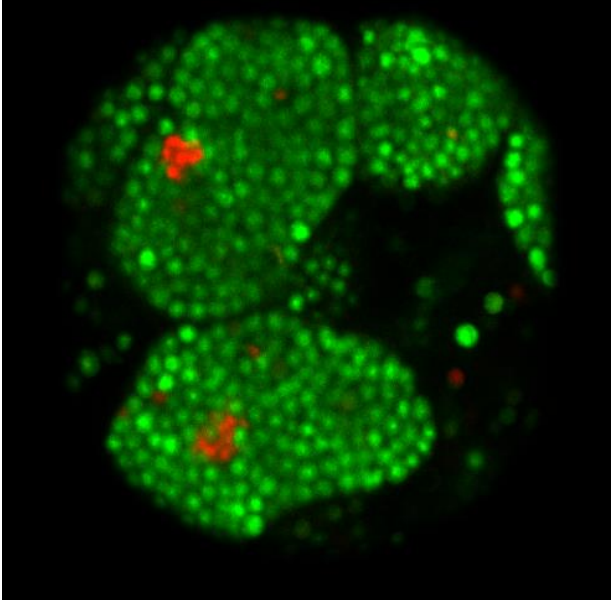
Wang, Q., Yang, S., Liu, J., Terecskei, K., Ábrahám, E., Gombár, A., *et al.* (2017) Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. *Proc Natl Acad Sci U S A* **114**: 201700715.

Wojciechowski, M.F., Sanderson, M.J., Steele, K.P., and Liston, A. (2000) Molecular phylogeny of the "Temperate Herbaceous Tribes" of papilionoid legumes: A supertree approach. *Adv Legum Syst* **9**: 277–298.

Yang, S., Wang, Q., Fedorova, E., Liu, J., Qin, Q., Zheng, Q., *et al.* (2017) Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proc Natl Acad Sci U S A* **114**: 201700460.

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.





Discussion

Partie III : Discussion

Observations générales concernant l'utilisation de données -omiques

Analyses différentielles en RNA-seq

Mon arrivée au laboratoire a coïncidé avec l'obtention du premier jeu de données de type RNA-seq. La quantification de l'expression d'un gène se fait cette fois par la cartographie des reads (lectures, fragments d'ADN séquencés) obtenus sur l'ensemble du génome de l'espèce étudiée. Contrairement aux données de microarrays provenant de mesures d'hybridation par fluorescence qui étaient des données continues, celles obtenues par RNA-seq sont discrètes. Elles ont donc nécessité la mise au point de nouveaux outils statistiques plus adaptés aux contraintes spécifiques du RNA-seq. Les modèles linéaires de type ANOVA proposés dans les packages R du type limma-voom n'étaient donc pas recommandés pour les données discrètes jusqu'à récemment (Ritchie *et al.*, 2015).

Malgré tout, trois problèmes liés à la technologie RNA-seq nécessitent également d'être pris en considération avant d'effectuer les analyses différentielles. Le premier est la variabilité de la profondeur de séquençage entre les différents échantillons qui vont être comparés, nécessitant une étape de normalisation basée sur la taille des différentes banques. Le deuxième est le contenu en GC des fragments d'ADN séquencés, dont l'occurrence de séquençage est plus faible quand les gènes sont riches en GC. Le troisième est l'effet de la longueur des gènes sur leur expression différentielle (Oshlack *et al.*, 2009). En effet, la plupart des protocoles de séquençage pour le RNA-seq passent par une approche de fragmentation des ADNc avant le séquençage pour gagner en profondeur de séquençage. Mais cela signifie que le nombre total de reads d'un gène est proportionnel à son niveau d'expression multipliée par sa taille. Donc plus un gène est grand plus le nombre reads qui lui seront associés sera important comparé à des gènes plus petits dont le niveau d'expression est similaire. Et sachant que la puissance statistique d'une expérience dépend de l'échantillonnage, la détection de l'expression différentielle sera plus puissante d'un point de vue statistique pour les gènes longs. Mais cela peut aussi créer une surreprésentation des fragments de ces gènes longs donnant de très fortes inductions pouvant être éloignées de la réalité.

Les premières méthodes de normalisation ont donc été basées sur le calcul des RPKM/FPKM (Reads/Fragments per Kilobase per Million reads/fragments) de tous les gènes, prenant ainsi en compte à la fois la profondeur de séquençage et la taille des gènes (Mortazavi *et al.*, 2008). C'est cette approche de normalisation qui avait été favorisée au sein du laboratoire pour les jeux de données RNA-seq de *Bradyrhizobium* sp. ORS285 et *B. diazoefficiens* USDA110. Cependant, il a été montré mathématiquement que cette méthode de normalisation ne corrige pas le biais lié à la longueur des gènes (Oshlack et Wakefield, 2009). Les méthodes de normalisation provenant des packages DeSeq2 et EdgeR ont été définies comme étant les plus convenables et nous ont été conseillées (Dillies *et al.*, 2012).

C'est pour cela qu'à la suite de mon premier comité de thèse, j'ai pris la décision de me former à ce type d'analyse en utilisant le package DEseq2 pour ces jeux de données. Ceci m'a permis d'être coauteur de deux articles supplémentaires au cours de ma thèse. (Gonzalez-Mula *et al.*, accepté dans New Phytologist ; Annexe 3 et Obayashi *et al.*, en révision).

Taille des gènes et expression des NRPS

A la suite de ces analyses transcriptomiques sur quatre jeux de données différents, j'ai remarqué que les gènes très longs de ces différentes bactéries qui m'ont mené à quelques interrogations. Comme mentionné et discuté dans le chapitre 3 sur les effets de la mutation du gène *bclA* sur la symbiose *Aeschynomene-Bradyrhizobium* sp. ORS285, les gènes appelés NRPS (Non Ribosomal Peptide Synthases) sont induits dans les conditions de vie libre et symbiotiques chez le mutant Δ*bclA*.

Ces gènes codent des protéines de très grande taille et composées de plusieurs modules, chacun impliqué dans l'élongation du peptide par un nouvel acide aminé, et donc capables de synthétiser des peptides différents de ceux obtenus via le processus majoritaire qu'est la traduction des ARN messagers (Hur *et al.*, 2012). Cette première observation chez ORS285 ne m'a pas étonné la première fois mais j'étais à la fois le premier de l'équipe à appréhender un jeu de données de type RNA-seq et n'avais pas d'autres expériences quant à leur manipulation.

Néanmoins, les très fortes inductions de ces gènes codant des NRPS vont bien dans le sens que plus un gène est grand, plus l'expression différentielle sera importante. J'ai observé le même phénomène chez *Agrobacterium* et chez *Burkholderia*. En conséquence je me demande

à quel point ces très fortes inductions correspondent une réalité biologique en termes de production des peptides non ribosomiques ?

La réponse à cette question est malheureusement relativement difficile à donner, car cela nécessiterait de créer un mutant pour ces clusters NRPS et d'être capable de détecter le métabolite produit entre le sauvage et le mutant dans des conditions induisant l'expression de ces NRPS. Cependant, l'expression de ces NRPS a été vérifiées chez la souche ORS285 par qPCR et a donné des résultats allant dans le même sens que les données RNA-seq (cf chapitre 2)

Approches protéomiques

Les approches protéomiques se sont beaucoup améliorées depuis une quelques années, passant de l'étude de gels 2D à la spectrométrie de masse. Une fois les échantillons préparés et digérés par la trypsine, les résultats provenant des analyses LC-MS/MS (*Liquid chromatography tandem mass spectrometry*) ont été effectuées par approche *shotgun*, c'est à dire sans décomplexifier les échantillons.

Les résultats de spectrométrie nous ont laissé la possibilité d'utiliser trois méthodes de quantification possibles des protéines présentes dans les différents échantillons. Les deux première méthodes sont le *peak counting* et le *spectral counting* qui sont des approches plutôt qualitatives informant sur les caractéristiques de présence/absence pour chaque protéine, mais ne permet pas une quantification fine de celles-ci. Au contraire la technique utilisant les XICs (pour *eXtracted Ion Current*) rendent possible une analyse quantitative des protéines, mais uniquement celles présentes dans l'ensemble des échantillons testés (Valot *et al.*, 2011, Langella *et al.*, 2013).

Etant donné que les conditions bactéroïdes et de vie libre sont très différentes, la méthode utilisant les XICs ne semble être plus la plus adaptée pour nos jeux de données car beaucoup de protéines ne seraient pas considérées. Nous nous sommes arrêtés sur l'approche basée sur le *spectral counting* car elle révèle la présence/absence de toutes les protéines tout en ayant une meilleure gamme dynamique que le *peak counting*.

Comme mentionné dans l'introduction, on peut observer que comparé aux approches transcriptomiques, l'identification et la quantification des protéines couvre une portion plus faible du génome. Les protéines membranaires ou les facteurs de transcriptions sont moins

faciles à extraire/identifier en protéomique. Mais l'interprétation faisant suite à ces observations sont plus fiables car on quantifie « ce qui peut faire arriver » comme les enzymes.

Nous avons observé de très bonnes corrélations transcriptomiques et protéomiques quand nous avons essayé d'intégrer les deux jeux de données pour la souche *B. diazoefficiens* USDA110 prélevés à des temps similaires post-inoculation (cf chapitre 4).

Pour finir, je souhaitais faire un parallèle entre les analyses statistiques de type RNA-seq et de *peak/spectral counting* en protéomique. Les valeurs testées sont dans les deux cas basées sur des données discrètes. Je m'interroge alors sur les raisons possibles de la non-utilisation de modèles similaires à ceux du RNA-seq (DEseq2) pour ces données protéomiques.

Aspects métabolomiques

Les analyses métabolomiques ont été réalisées sur des nodosités complètes. Ces expériences ont permis d''avoir une idée du fonctionnement général d'une nodosité de *A. afraspera*, *A. indica* et *G. max* en fonction de leurs symbiontes compatibles respectifs. Nous avons notamment pu distinguer quels sont les métabolites produits et exportés des nodosités comme source d'azote pour le reste de l'hôte végétal.

S'il était déjà connu que le soja exporte l'azote fixé sous forme d'uréides (allantoine), nos analyses montrent que l'azote est exporté sous forme d'asparagine chez *A. indica* et d'un mélange d'asparagine et de glutamine chez *A. afraspera*. Les résultats les plus intéressants proviennent de la comparaison entre les nodosités induites par une bactérie sauvage et une autre bactérie mutante incapable d'effectuer une symbiose fonctionnelle. Cette comparaison vise à possiblement identifier des métabolites habituellement consommés par les bactéroïdes fixateurs d'azote du fait de leur accumulation lorsque cette même fixation n'est pas active.

Néanmoins, l'utilisation de nodosités complètes ne permet pas de savoir quelle est la part de l'hôte végétal ou des microsymbiontes dans l'accumulation de la plupart des métabolites C'est également une vision figée alors que les métabolites circulants pourraient être détectés par des analyses de fluxomique.

Pour aller plus loin dans l'exploration du métabolisme des bactéroïdes différenciés, il faudrait parvenir à extraire les bactéroïdes de leurs cellules-hôtes végétales. Un protocole d'extraction a été testé au laboratoire au cours de ma thèse. Mais il reste à mesurer la pureté de

l'échantillon obtenu via ce protocole, car la contamination par des organites végétaux est fort probable.

é hodes d'in é ra ion des echnolo i es -omiques

En général, chaque technologie (transcriptome, protéome, métabolome, etc.) est analysée indépendamment par le biais de méthodes statistiques univariées, notamment ANOVA, modèles linéaires ou tests t. Cependant, un tel type d'analyse ignore les relations entre les différents objets d'étude de chaque technologie, et peut manquer des informations biologiques cruciales. Les approches multivariées, qui modélisent les caractéristiques du modèle en tant qu'ensemble, peuvent donc fournir une image plus perspicace d'un système biologique et compléter les résultats obtenus à l'aide de méthodes univariées. Ces analyses multivariées sont très puissantes quand on intègre des données -omiques qui sont appariées, notamment en prenant en utilisant l'infrastructure logicielle appelée DIABLO disponible dans le package R mixOmics (Rohart *et al.*, 2017).

Ce n'est malheureusement pas sur les mêmes échantillons qu'ont été effectuées les différentes analyses transcriptomiques, protéomiques et métabolomiques présentées dans ce manuscrit. Cela n'empêche pas d'être capable d'intégrer « à la main » les informations congruentes entre les différentes technologies pour afficher un meilleur degré de certitude quant aux phénomènes observés. Ceci a déjà été réalisé sur *B. diazoefficiens* USDA110 avec un transcriptome et un protéome générés séparément mais tout de même intégrés (Pessi *et al.*, 2007 ; Delmotte *et al.*, 2010).

Ces niveaux d'intégration des données -omiques vont sans l'ombre d'un doute se généraliser et probablement être utilisés en routine dans beaucoup de travaux scientifiques futurs. Il sera important de garder à l'esprit la nécessité de travailler sur des données appariées.

Risques associés aux approches sans a priori comme point de départ pour découvrir le rôle de nouveaux gènes

Les analyses fonctionnelles effectuées sur les souches ORS285 et USDA110 basées sur l'expression différentielle des gènes par rapport à la culture libre ou entre plantes hôtes ont été infructueuses. Au cours de cette thèse, 28 mutants d'insertion et 6 de délétion ont été générés sur la souche ORS285, ainsi que 10 mutants d'insertion et 4 de délétion chez USDA110. Ces résultats mènent à la remise en question du paradigme répandu corrélant expression génique et

nécessité d'un gène dans une condition physiologique donnée. Il y a bien entendu beaucoup de paramètres à prendre en compte pouvant expliquer ces observations.

Redondance génétique/fonctionnelle

Tout d'abord, les génomes de rhizobia font partie des plus grands génomes bactériens (entre et 5 et 10 Mb) avec pour certaines espèces de nombreux plasmides symbiotiques ou impliqués dans la vie libre (Poole *et al.*, 2018). *Bradyrhizobium diazoefficiens* USDA110 possède par exemple plus 9000 gènes sur son chromosome. Il n'est donc à première vue pas étonnant d'y retrouver plusieurs gènes jouant un rôle similaire dans le génome. Mais cette affirmation est loin d'être évidente car la redondance fonctionnelle peut rapidement impliquer un relâchement de la sélection naturelle et la pseudogénisation, et donc être instable. Cependant, de nombreux modèles théoriques ont montré des possibilités de redondance fonctionnelles stables évolutivement qui soient applicables pour des procaryotes haploïdes (Nowak *et al.*, 1997).

Les redondances génétiques peuvent même être considérées comme des circuits de secours réactifs. C'est le cas des gènes *Fsk1* et *Fsk2* de *Saccharomyces cerevisiae* impliqués dans la synthèse de 1,3-β-glucane. Les deux gènes sont exprimés de façon quasi similaire, mais l'un d'eux voit son expression augmentée si l'autre n'est plus fonctionnel (Kafri *et al.*, 2006). Pour les gènes candidats sélectionnés et mutés, il est très possible qu'il existe une forme de redondance fonctionnelle. Nous avons fait en sorte d'éviter au maximum cette dernière en comparant les degrés d'identité possibles des candidats avec les autres gènes de l'organisme. S'il existe des gènes très similaires et que la mutation d'un seul d'entre eux n'altère pas le fonctionnement de la symbiose, il conviendrait de faire des séries de mutations sur ces gènes pour pouvoir conclure sur le rôle de ces gènes en conditions symbiotiques ou même en vie libre.

Pour aller plus loin, il serait aussi intéressant de comparer l'expression des gènes en question en conditions d'induction pour les différents mutants possibles générés. Cela permettrait d'observer ou non des compensations d'expression du gène restant quand l'(es) autre(s) gène(s) est/sont inactivé(s).

es hau s niveau d'e ression des ènes ne son as corrélés à leur nécessi é dans une condition donnée

Il est possible que les modifications transcriptionnelles ou traductionnelles observées entre les différentes conditions ne soient pas des réponses adaptées à la condition symbiotique, en termes d'unicité (induit uniquement en conditions symbiotiques) et de nécessité (requis pour une symbiose fonctionnelle). Les comparaisons systématiques entre la valeur sélective de l'inactivation d'un gène et son expression ont montré que ces deux paramètres sont très peu corrélés. Une première étude sur la levure a montré que seul 7% des gènes qui montrent une augmentation dépendante de la condition biologique testée sont effectivement requis pour celle-ci (Giaever *et al.*, 2002).

L'émergence des technologies du type Transposon-sequencing (Tn-seq) combinées à des études transcriptomiques montrent que les réponses transcriptionnelles sont en grande partie suboptimales. C'est le cas chez la bactérie *Shewanella oneidensis* MR-1, qui présente des coefficients de corrélation entre valeurs sélectives relatives et expression relatives proches de zéro dans différentes conditions comparées (Deutschbauer *et al.*, 2011 ; Price *et al.*, 2013). La même remarque peut être faite en considérant l'étude du pathogène biotrophe *Agrobacterium tumefaciens* C58, qui induit la formation de tumeurs chez les plantes en exploitant de nombreux métabolites (Gonzalez-Mula *et al.*, accepté dans New Phytologist ; Annexe 3).

Chez *Agrobacterium*, aucun des gènes différentiellement exprimés en comparant les transcriptomes réalisés lorsque la bactérie croit avec du GABA ou du saccharose comme source de carbone n'est essentiel à l'assimilation du GABA. Les gènes *blcAB* (ou *attKL*) sont les plus induits avec le GHB ou le GABA comme source de carbone par rapport à la référence, mais *blcA* et *blcB* ne sont requis que pour croître en présence de GHB (Carlier *et al.*, 2004 ; Chai *et al.*, 2007). Ces exemples renforcent l'idée que l'expression d'un gène n'est pas liée à sa nécessité avec des effets visibles au niveau phénotypique. L'expression des gènes n'est pas toujours finement contrôlée par rapport à la condition. Il est souvent considéré à tort que les bactéries sont des systèmes optimisés par des millions d'années d'évolution. Mais la spécificité des mécanismes de régulation de l'expression peut être relativement lâche et entrainer des expressions de gènes « futiles ». En reprenant l'exemple précédent, l'expression non nécessaire des gènes *blcAB* en présence de GABA pourrait être liée à la proximité chimique entre le GABA et le GHB qui lèveraient tous deux de la même façon l'inhibition de l'expression de *blcAB* en se fixant au répresseur AttJ (Chai *et al.*, 2007).

Limites à la découverte de phénotype liée aux conditions de laboratoire

Un autre paramètre non pris en compte au cours de mon travail de thèse est la spécificité des conditions de laboratoire dans lesquelles sont effectuées les analyses fonctionnelles en routine depuis des années. En effet, l'inoculation des bactéries se fait en général avec une charge bactérienne très importante. En fonction de l'association étudiée, on verse entre 2 et 10mL de suspension bactérienne à 10^8 bactéries/mL pour chaque plante. Sachant qu'un gramme de sol contient 10^4 espèces de bactéries différentes et 10^9 bactéries et si l'on suppose une répartition uniforme en nombre de ces espèces bactériennes, les rhizobia sont mille fois plus concentrée dans nos conditions de laboratoire (Poole *et al.*, 2018). Cette méthode n'empêche pas la découverte de gènes dont l'inactivation compromet la capacité de la bactérie à réaliser une symbiose fonctionnelle avec son hôte.

Cependant, on s'affranchit sans doute de certaines contraintes comme l'aptitude à coloniser efficacement la racine. Un exemple d'intérêt de notre jeu de données est le gène BRAD285_v2_1078. Ce dernier est probablement lié à l'environnement de la plante ou à la persistance dans la rhizosphère. Il code une protéine homologue des protéines de surface *S-layer*. Elles sont très diverses parmi les espèces bactériennes (Fagan et Fairweather, 2014). Ces protéines *S-layer* sont plausiblement sécrétées via le système de sécrétion de type 2 sont les plus à même d'être en contact avec des hôtes eucaryotes potentiels. Cela implique une interaction possible avec le système immunitaire de l'hôte et les molécules de surface de l'hôte. Nous avons d'abord imaginé un rôle potentiel dans la résistance aux peptides NCR car la S-layer de *Caulobacter crescentus* est impliquée dans la résistance aux peptides antimicrobiens (de la Fuente- Nuñez *et al.*, 2012). L'absence de phénotype symbiotique nous mène aujourd'hui à supposer un rôle dans l'attachement au système racinaire des *Aeschynomene*, comme déjà observé chez *Azotobacter vinelandii* (Liew *et al.*, 2015). Mais les conditions de laboratoire utilisées font que l'inoculation se fait avec une telle concentration en bactéries qu'il est fort probable qu'on ne puisse pas observer un phénotype plus fin qui pourrait exister. Le locus *gsp1-gsp2* dont la large délétion ne présente pas de phénotype symbiotique visible dans les conditions habituelles (et comprenant le gène BRAD285_v2_1078) est à tester en priorité pour augmenter les chances d'observer un phénotype de colonisation (Lamouche *et al.*, 2018).

Il conviendrait alors de chercher à mesurer l'efficacité de l'infection en fonction de la quantité de bactéries inoculée pour une plante. Le but serait de partir de la condition classique de laboratoire et faire des dilutions en séries de l'inoculum jusqu'à obtenir une concentration limite à partir de laquelle on n'observe plus l'apparition de nodosités. Cette concentration limite

serait alors à comparer entre les mutants bactériens générés et la souche sauvage pour estimer la capacité des souches à s'attacher à la racine. Pour aller plus loin, il serait pertinent de mesurer la compétitivité de colonisation rhizosphérique des souches mutantes par rapport à la souche sauvage par co-inoculation à une concentration proche de la concentration limite de la souche sauvage estimée au préalable. De même, mesurer cette compétition à une concentration non limitante (10 fois plus concentrée que la concentration limite par exemple) permettrait de mettre en compétition les deux souches pour la colonisation des nodosités.

Enfin, un article récent utilisant l'approche Tn-seq pour détecter les gènes bactériens de *Pseudomonas simiae* WCS417r ayant une perte de valeur sélective concernant l'attachement au rhizoplan d'*Arabidopsis thaliana* Col-0 a été publié (Cole *et al.*, 2017). À la lumière de ces résultats, nous pourrions chercher des orthologues de ces gènes chez ORS285/USDA110 pour sélectionner des gènes candidats impliqués dans l'attachement aux racines d'*Aeschynomene* après avoir vérifié leur expression. Il en ressort sans surprise de nombreux gènes codant des sous-unités du flagelle bactérien qu'il conviendrait de muter et phénotyper en conditions symbiotiques, ainsi que d'autres protéines non testées auparavant. Les résultats très intéressants de cette étude justifient également de faire des expériences similaires avec des banques Tn-seq de *E. meliloti* 1021 colonisant *M. sativa*, ainsi que *Bradyrhizobium* sp. ORS285 colonisant *A. afraspera* et *A. indica*.

Le processus de différenciation des bactéroïdes au cours des interactions symbiotiques entre *Bradyrhizobium* et *Aeschynomene*

La différenciation en bactéroïdes sphériques, un avantage sélectif pour le partenaire végétal ?

Nos diverses études ont tenté de confronter l'existence possible d'une corrélation entre le processus de différenciation des bactéroïdes et le bénéfice tiré par le partenaire végétal. Les résultats suggèrent des observations qui peuvent être à première vue relativement contradictoires. Notre première étude sur les associations *Ensifer-Medicago* semble indiquer l'existence d'une corrélation positive entre efficacité symbiotique et différenciation, impliquant l'augmentation de taille et de ploïdie des bactéroïdes (Kazmierczak *et al.*, 2017 ; Annexe 1). Cependant, cette observation ne se retrouve pas en tout point dans l'étude des associations *Bradyrhizobium-Aeschynomene*. En effet dans ce cas, la corrélation entre efficacité et différenciation semble plus diffuse car elle est positive et significative quand comparée à la taille des bactéroïdes, mais nulle quand comparée au niveau de ploïdie de ces derniers.

Cependant, nous avons tout de même pu observer une augmentation de l'efficacité symbiotique quand les bradyrhizobia sont associées aux *Aeschynomene* imposant un programme de différenciation plus poussé. En effet, les *Aeschynomene* portant des bactéroïdes sphériques (*A. indica* et *A. evenia*) semblent être plus efficaces que celles imposant une différenciation en bactéroïdes « allongés » (*A. afraspera* et *A. nilotica*) (Czernic *et al.*, 2015, Lamouche *et al.*, 2018, chapitres 1 et 2). D'autre part, il semblerait que chez les IRLC, le niveau de différenciation des bactéroïdes est positivement corrélé à la taille du répertoire de NCR cationiques (Montiel *et al.*, 2017). L'expression des gènes codant les NCR présente un coût important pour l'hôte végétal. En effet chez *M. truncatula*, 10% des transcrits d'une nodosité sont des ARNm de peptides NCR (Roux *et al.*, 2014). La conservation d'un tel mécanisme coûteux pour la plante, qui plus est apparu plusieurs fois de manière convergente, semble indiquer un avantage certain pour la plante en termes de rendement sur le processus symbiotique.

Les données génomiques des Dalbergioïdes sont relativement limitées, mais il est possible de spéculer que le répertoire de peptides NCR-like diffère quantitativement et qualitativement entre les différentes espèces d'*Aeschynomene* induisant des bactéroïdes sphériques ou allongés (Czernic *et al.*, 2015). Il est en effet possible que le cocktail de NCR-like sécrétés dans les nodosités ayant des bactéroïdes sphériques exerce une action plus contraignante au niveau du cycle cellulaire et du métabolisme des symbiontes. Cependant, il y a fort à parier que même si la différenciation médiée par les peptides NCR affecte l'efficacité du processus symbiotique, d'autres facteurs inconnus peuvent également jouer un rôle dans la modulation de cette efficacité symbiotique chez les Dalbergioïdes. Par ailleurs, des données transcriptomiques d'une autre plante de la famille des Dalbergioïdes, *Arachis hypogaea,* ont récemment été publiées, mettant en lumière l'existence de peptides dérivant de la protéine de défense PR-1 (PR : Pathogen-Related) appelés CAPE (Cationic Antimicrobial Peptide Derived Peptide) (Karmakar *et al.*, 2018). *A. hypogaea* impose également un processus de différenciation en bactéroïdes sphériques (Sen et Weaver, 1984). Aucun peptide NCR-like n'a cependant été détecté dans ce jeu de données, posant alors la question du rôle de ces peptides CAPE. Sont-ils impliqués dans le processus de différenciation ? Et si oui, est-ce le résultat d'une convergence d'utilisation de peptides antimicrobiens comme effecteurs de la différenciation des bactéroïdes, remplaçant ainsi le rôle des NCR ? Ou bien leur fonction serait simplement liée à un/des autre(s) phénomène(s) ?

Les occurrences en faveur de la première possibilité ne manquent pas dans le vivant. Il existe pléthore de systèmes symbiotiques entre des hôtes Eucaryotes et symbiontes Procaryotes où les hôtes utilisent des peptides antimicrobiens en dehors de la symbiose rhizobium-légumineuses. C'est notamment le cas de symbioses entre des insectes et des bactéries symbiotiques, dont celle entre *R. pedestris* et *Burkholderia* sp. RPE64 où des peptides antimicrobiens appelés CCR (Crypt-specific Cystein Rich) sont synthétisés au niveau de l'organe symbiotique (Futahashi *et al.*, 2013 ; Mergaert *et al.*, 2017). Également dans le domaine de la fixation d'azote, un peptide d'*Alnus glutinosa*, Ag5, a été découvert comme exprimé spécifiquement dans les nodosités actinorhiziennes infectées par les bactéries du genre *Frankia* (Carro *et al.*, 2015). Ag5 semble être un peptide antimicrobien impliqué dans la perméabilisation des membranes permettant le relargage de métabolites azotés comme la glutamine et le glutamate, alors disponibles pour la plante hôte (Mergaert *et al.*, 2017). C'est en tout cas une hypothèse intéressante que de considérer l'effet peptides comme facilitateurs de l'échange nutritionnel entre l'hôte et son symbionte, bien que les effets puissent être délétères pour ce dernier.

Un processus de différenciation à reconfirmer pour les bactéroïdes de A. afraspera

Au cours de ce travail de thèse, j'ai pu observer et confirmer maintes fois le processus de différenciation des bradyrhizobia quand associées aux plantes du groupe d'inoculation III telles que *A. evenia* et *A. indica* (cf chapitres 1 et 2), comme publié avant ma thèse (Czernic *et al.*, 2015, Figure 52A). La même observation n'a cependant pas été confirmée pour les bactéroïdes de *A. afraspera*, appartenant au groupe d'inoculation II, malgré les répétitions expérimentales (Figure 52B).

Les mesures de ploïdie par cytométrie en flux que j'ai effectuées utilisant de l'iodure de propidium ont donné des niveaux de ploïdie beaucoup plus résolutifs, avec des pics très bien définis avec les bradyrhizobia contrairement au DAPI utilisé jusqu'alors comme dans Mergaert *et al.*, (2006) et Czernic *et al.*, (2015). Les valeurs de ploïdie observées pour les bactéroïdes des trois espèces d'*Aeschynomene* mentionnées sont plus faibles que ce qui a été publié. En fonction de la souche utilisée, les bactéroïdes sphériques se différencient bel et bien, mais avec une valeur de ploïdie médiane entre 3C et 6C. Les bactéroïdes « allongés » de *A. afraspera* ont une ploïdie qui se situe entre 1C et 2C, comme la culture libre de bactéries (cf chapitre 1). Ces résultats ne sont donc pas en accord avec ce qui a été publié précédemment. Il serait convenable

de tester à nouveau l'observation de bactéries libres et de bactéroïdes extraits de nodosités après marquage à l'iodure de propidium pour l'ensemble des associations utilisées. L'effet des peptides NCR-like de *A. afraspera* pourrait en réalité se limiter à perméabiliser l'enveloppe bactérienne, de la même façon de manière que le peptide Ag5 qui perméabilise les vésicules de *Frankia* lorsqu'associées à *Alnus* (Carro *et al.*, 2015). Ce processus de différenciation serait plus « doux » que celui imposé par *A. indica* ou *A. evenia*, sans modifier drastiquement le cycle cellulaire ni la morphologie du symbionte.

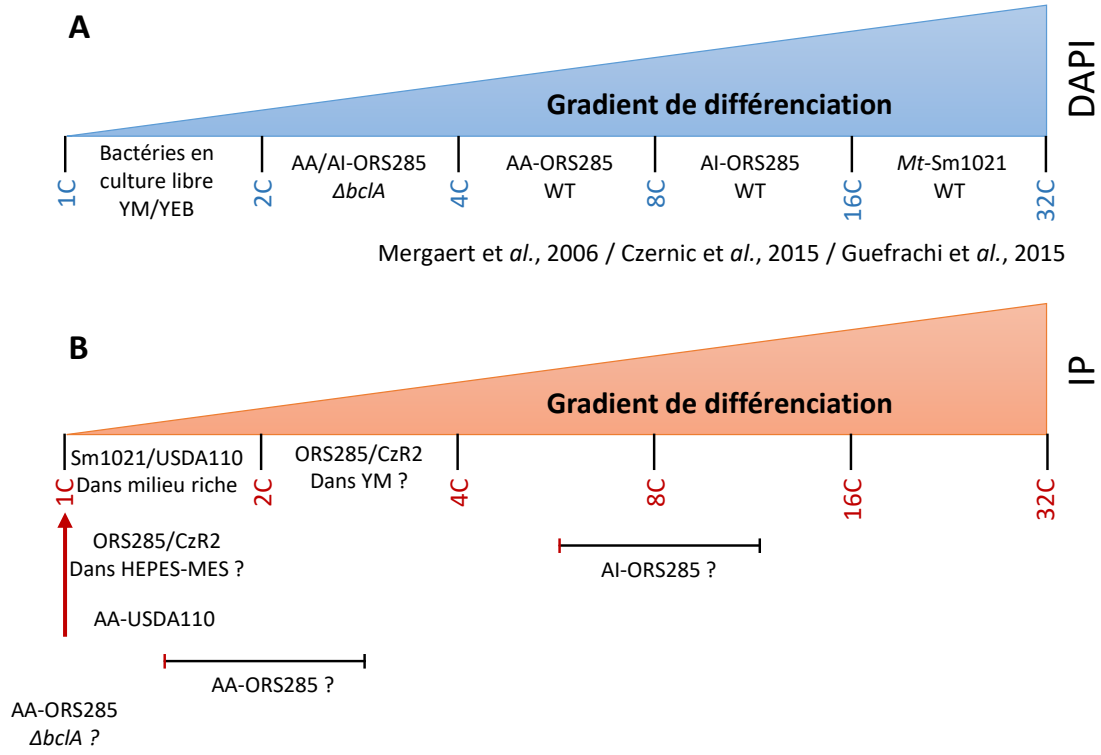


Figure 52: Modèles intégrant la diversité et l'intensité des processus de différenciation des bactéroïdes existant chez les IRLC et les Dalbergioïdes, basée sur une échelle de niveaux de ploïdie. (A) Modèle basé sur les publications mesurant les profils de ploïdie de différentes espèces bactériennes en culture libre ou bactéroïdes en utilisant le DAPI. (B) Modèle possible des processus de différenciation issu de mes résultats de thèse et d'autres non publiés de l'équipe. Ce modèle reste à clarifier concernant la symbiose *Bradyrhizobium-Aeschynomene* et utilisant l'iodure de propidium (PI) comme fluorophore. Les valeurs médianes actuellement mesurées pour les bactéroïdes de *A. afraspera* (AA) et *A. indica* (AI) sont celles de la borne inférieure (en rouge), mais pourraient être redéfinies si une nouvelle condition de calibration voit le jour. Mt : *Medicago truncatula*, ORS285 : *Bradyrhizobium* sp. ORS285 ; USDA110 : *B. diazoefficiens* USDA110 ; Sm1021 : *E. meliloti* 1021.

Mais le problème peut aussi venir de la culture de référence. En effet, les bradyrhizobia semblent sujettes à des changements morphologiques en fonction des sources de carbone qui leur sont données. *Bradyrhizobium* sp. 32H1 tend à s'allonger ou à devenir sphérique quand cultivé en milieu enrichi en arabinose et succinate, respectivement (Ramaswamy et Bal, 1986). Ces métabolites semblent d'ailleurs induire induire la fixation de l'azote en vie libre pour cette

bactérie. Plus récemment, une souche proche de *Bradyrhizobium arachidis* nommée CzR2 présente des changements morphologiques dépendant du milieu de culture (Huang *et al.*, 2016).

La morphologie de cette bactérie est celle d'un bacille classique dans du milieu HEPES+MES, mais est particulièrement allongée en présence d'extraits de levure et de mannitol (Figure 53). Il se trouve que ces deux composés sont les principales sources nutritives du milieu YM utilisé tout au long de cette thèse pour cultiver les bradyrhizobia. Ce milieu a été considéré comme celui de référence également pour les analyses -omiques et les analyses de cytométrie en flux. Il est donc possible que cela soit cette condition de référence YM qui fausse les observations dudit processus de différenciation chez *A. afraspera*. Refaire des expériences en cytométrie en flux pour vérifier l'allongement effectif des bradyrhizobia photosynthétiques en milieu YM par rapport à un milieu minimum permettra sans doute de mieux calibrer la condition de référence, tant en termes de taille des cellules que de ploïdie (Figure 52B). Il sera alors possible de (re)-confirmer la possible différenciation des bactéroïdes d' *A. afraspera* et de rediscuter de la différenciation « atypique » de la souche *B. diazoefficiens* USDA110 qui ne semble pas affectée morphologiquement par le milieu de culture YM comme c'est le cas pour la souche CzR2 et pourrait l'être pour ORS285 (Huang *et al.*, 2016).

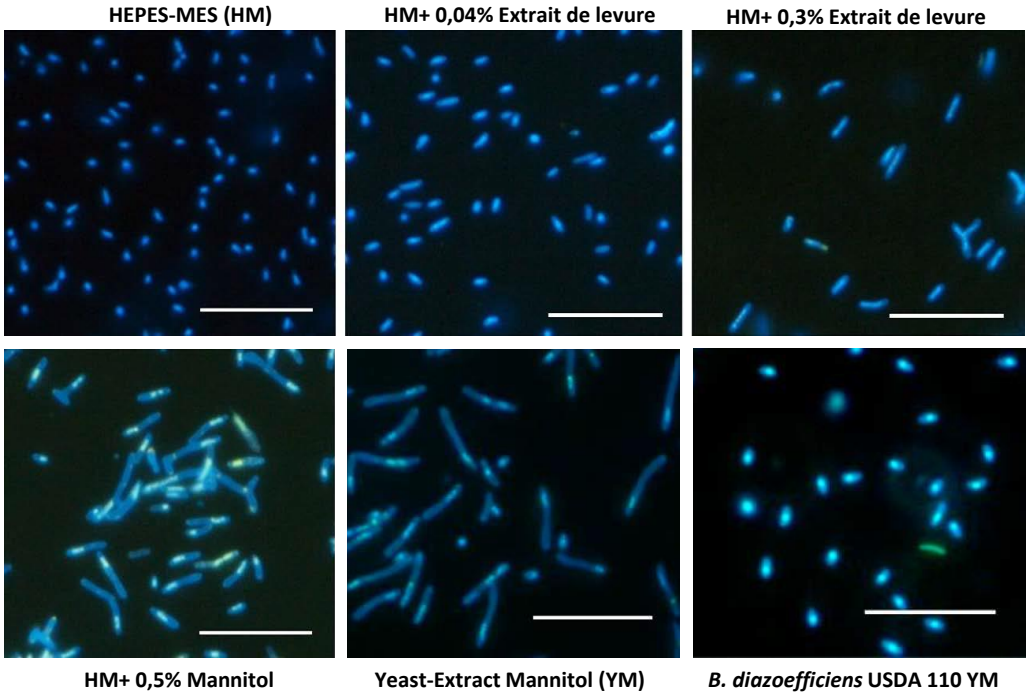


Figure 53: Comparaisons morphologiques de culture de la souche CzR2 dans différents milieux de culture. La présence d'extraits de levure et de mannitol agrandit considérablement la taille des bactéries de cette souche, mais n'a pas d'effet sur la souche *B. diazoefficiens* USDA110 (d'après Huang *et al.*, 2016).

’im or ance du e i do l c ane dans les associa ions Bradyrhizobium-Aeschynomene

La souche *B. diazoefficiens* USDA110 est un symbionte naturel du soja, mais nodule également *A. afraspera.* A la différence de la souche ORS285, *B. diazoefficiens* USDA110 ne nécessite pas la présence de la protéine BclA pour effectuer une symbiose fonctionnelle avec *A. afraspera* (Barrière *et al.*, 2017 ; Annexe 2). Si on admet que c’est la protéine BclA qui est requise pour faire face à l’effet délétère des peptides NCR-like, il semblerait que la souche USDA110 soit naturellement résistante à leurs effets. Cette résistance accrue aux peptides NCR-like pourrait s’expliquer par des propriétés de son enveloppe, affectant le peptidoglycane, les hopanoïdes et/ou le lipopolysaccharide.

Un premier candidat pourrait être le peptidoglycane. Il a été observé que lorsque la formation de celui-ci est perturbée, on observe des associations symbiotiques non fonctionnelles entre différentes espèces d’*Aeschynomene* et les souches *Bradyrhizobium* sp. ORS278, ORS285 ainsi que *B. diazoefficiens* USDA110 compatibles (Gully *et al.*, 2016). Cette découverte a été permise via l’analyse des phénotypes symbiotiques d’un mutant pour un gène codant une DD-carboxypeptidase. Son rôle dans la symbiose a été découvert via un crible génétique par insertion de transposon chez la souche ORS278 associée à *A. indica* (Bonaldi *et al.*, 2010). Les phénotypes symbiotiques observés sont une altération de la morphologie des bactéroïdes, ainsi qu’une augmentation du niveau de ploïdie par rapport à la souche sauvage (Gully *et al.*, 2016). La fonction catalytique de cette protéine est de cliver le domaine pentapeptidique des précurseurs de peptidoglycane, réduisant alors le maillage entre les différentes couches composant le peptidoglycane (Typas *et al.*, 2011). Cela permet un remodelage de ce dernier que l’on suppose nécessaire au processus symbiotique dans le cadre des divisions cellulaires et de ainsi que le transfert de composés de surface. Ces derniers pourraient être impliquées dans la limitation des réactions de défense de la plante comme c’est le cas du lipopolysaccharide ou des exopolysaccharides (Gourion *et al.*, 2015 ; Gully *et al.*, 2016 ; Zipfel et Oldroyd, 2017). Le mutant DD-carboxypeptidase des souches de *Bradyrhizobium* est plus sensible au stress salin et possède des morphologies extrêmes en symbiose avec les *Aeschynomene*. La souche mutée de USDA110 présente une morphologie anormale et est incapable de fixer l’azote quand elle est associée à *A. afraspera*. Ce phénotype n’est en revanche pas retrouvé lors de l’association avec le soja (Gully *et al.*, 2016). Pour aller plus loin, nous avons construit un double mutant pour les gènes *bclA* et DD-carboxypeptidase afin de mesurer leurs potentiels effets additifs et/ou épistatiques sur le processus symbiotique. Nous avons finalement observé que les deux gènes agissent indépendamment sur le processus

de différenciation, et que le phénotype symbiotique observé avec *A. afraspera* est similaire à celui de la mutation du gène de la DD-carboxypeptidase seule chez *B. diazoefficiens* USDA110 (Barrière *et al.*, 2017 ; Annexe 2). Malheureusement, les mêmes observations peuvent être faites avec ORS285 et ORS278 avec leurs plantes hôtes (Gully *et al.*, 2016 ; Barrière *et al.*, 2017 ; Annexe 2). L'importance du remodelage du peptidoglycane semble donc ne pas être une spécificité de la souche USDA110 pouvant expliquer l'effet limité des peptides NCR-like quand elle est associée à *A. afraspera*, mais peut tout de même être un paramètre important parmi d'autres dans le processus de résistance qui est probablement multifactoriel.

Les hopanoïdes, des composés lipidiques aux rôles multiples

Un autre composé d'intérêt pourrait être lipidique, il s'agit des hopanoïdes. Ce sont des triterpènes pentacycliques similaires au cholestérol des cellules eucaryotes et jouant des rôles similaires dans la stabilisation et la fluidité membranaires (Belin *et al.*, 2018). Ils sont présents à la fois dans les membranes interne et externe des bactéries. En plus de leur rôle déjà mentionné dans la limitation de la pO_2 chez *Frankia* nécessaire à la fixation d'azote (Berry *et al.*, 1993 ; Kleeman *et al.*, 1994 ; Ghodhbane-Gtari *et al.*, 2014), il est fort possible que la résistance de USDA110 aux peptides antimicrobiens provienne de sa composition lipidique particulière comportant jusqu'à 40% d'hopanoïdes (Kennenberg *et al.*, 1995). En lien, la mutation du gène *shc* codant une squalène-hopène cyclase (Shc), une enzyme clé de la voie de biosynthèse des hopanoïdes, est létale chez cette bactérie (Kulkarni *et al.*, 2015). La mutation de ce gène chez *Bradyrhizobium* sp. BTAi1 a pourtant été un réalisée, suggérant que les hopanoïdes jouent un rôle prépondérant chez la souche *B. diazoefficiens* USDA110 dans la résistance aux stress environnementaux (Silipo *et al.*, 2014). Le squelette de base des hopanoïdes synthétisé par la Shc peut ensuite être modifié par d'autres enzymes (Belin *et al.*, 2018). Deux mutants affectés dans la modification de ces molécules ont été produits chez USDA110. Le gène *hpnH* est reponsable de l'élongation du squelette carboné en C_{30} en adénosylhopane, passant de C_{30} à C_{35} par l'ajout d'une adénosine. Le gène *hpnP* est impliqué dans la méthylation de différent hopanoïdes formés en aval de l'effet de HpnH (Belin *et al.*, 2018). Les résultats majeurs de cette étude sont la diminution de la résistance au peptide NCR335 de *Medicago* du mutant Δ*hpnH* qui n'est par ailleurs ni capable d'effectuer une symbiose fonctionnelle avec *A. afraspera*, ni d'y maintenir une infection chronique (Kulkarni *et al.*, 2015). De manière intéressante, ce même mutant n'est pas affecté lors de la symbiose avec le soja, et l'une des différences majeures entre ces deux conditions est la production de peptides NCR-like chez *A.*

afraspera. Cela suggère alors la possibilité d'une corrélation entre présence de ces hopanoïdes et la résistance aux peptides NCR-like d'*A. afraspera*. Cependant, l'effet de ces peptides NCR peut être discuté du fait de la morphologie des nodosités d'*A. afraspera* infectées par le mutant Δ*hpnH*. Ces dernières présentent quelques phénotypes de réactions de défense avec une colonisation très variable, ne ressemblant pas aux nodosités des mutants Δ*bclA* de la souche ORS285 entièrement colonisés mais où les bactéries meurent dans les symbiosomes sans présenter de réactions de défense du côté végétal (Guefrachi *et al.*, 2015). Les hopanoïdes semblent être une piste majeure pour expliquer l'absence quasi-totale de différenciation de *B. diazoefficiens* USDA110.

Le lipopolysaccharide, un carrefour de signalisation des interactions Eucaryotes-Procaryotes

Le LPS est le composé le plus externe de l'enveloppe des bactéries à coloration de Gram négative. La structure du LPS est très spécifique de la souche bactérienne, et il est globalement constitué de trois domaines. Il est ancré à la membrane par sa partie hydrophobe appelé lipide A. Il y a ensuite un cœur oligosaccharidique qui fait le lien entre le lipide A et un polysaccharide appelé antigène O spécifique, très variable d'une souche à une autre et lui conférant une part de son immunogénicité. Chez *B. diazoefficiens*, les gènes tels que *galE* et *rfaL* impliqués dans la synthèse et le greffage de l'antigène O sur le LPS sont requis pour l'établissement de la symbiose car leurs mutants respectifs ont un phénotype Nod$^-$ lors de la symbiose avec le soja (Noh *et al.*, 2015 ; Chang *et al.*, 2015). Il semblerait également que l'absence de la partie oligosaccharidique du LPS mène à l'arrêt du développement des cordons d'infection chez *Rhizobium leguminosarum* bv. *trifolii* (Carlson *et al.*, 1986). Ces données suggèrent que l'antigène O joue un rôle important lors de nombreuses associations symbiotiques, mais aussi dans la reconnaissance entre plantes et pathogènes (Zipfel *et al.*, 2017). Cependant, les gènes homologues des gènes *rfaL* nécessaires pour la formation de l'antigène O n'ont pas de rôle chez ORS285 pour infecter les *Aeschynomene* et former une symbiose fonctionnelle (Busset *et al.*, 2016).

Le LPS des bradyrhizobia photosynthétiques (comme ORS285) présente néanmoins quelques spécificités, notamment leur antigène O est formé par un unique sucre spécifique, le bradyrhizose, qui est également non immunogénique (Silipo *et al.*, 2011). *B. diazoefficiens* a quant à lui en antigène O classique non composé de bradyrhizose. Ces différences phénotypiques observés entre *Aeschynomene* et soja soulève une interrogation. Quel pourrait

être le phénotype symbiotique du mutant *rfaL* (ou autre gène impliqué dans la synthèse du O-antigen) de *B. diazoefficiens* en association avec *A. afraspera* ? Si celui-ci est également Nod$^-$, cela impliquerait des différences entre les deux espèces de *Bradyrhizobium*, voire même entre les bradyrhizobia photosynthétiques et non photosynthétiques, quant à leur capacité à inhiber les défenses immunitaires de leur partenaire végétal pour la mise en place de l'interaction symbiotique. On peut en effet considérer que les mutations de l'antigène O peuvent révéler d'autres motifs moléculaires capables d'éliciter des réactions de défense de la part de la plante. Il est également possible que *B. diazoefficiens* USDA110 module la structure de son antigène-O lorsqu'au contact de peptides NCR. Cela a déjà été montré pour *Ensifer fredii* HH103 élicitant des nodosités avec *Glycyrrhiza uralensis*, une IRLC productrice de peptides NCR (Crespo-Rivas *et al.*, 2016 ; Montiel *et al.*, 2017). Si la symbiose n'est pas perturbée, alors on pourrait émettre une l'hypothèse que c'est possiblement le processus de colonisation initié par cordons d'infection qui requiert l'existence de l'antigène O du LPS pour la mise en place de l'interaction, là où l'infection par « crack-entry » ne les nécessite pas comme observé chez ORS285 (Busset *et al.*, 2017).

La particularité du lipide A des bradyrhizobia comme, *B. diazoefficiens* USDA110, est son association à deux VLCFA (Very Long Chain Fatty Acid). De nombreuses bactéries n'en possèdent pas et les autres rhizobia n'en possèdent qu'un seul sur leur lipide A. Il a été suggéré que les que les VLCFA liés au lipide A facilitent un mode de vie intracellulaire en augmentant la stabilité membranaire, car également retrouvé chez des pathogènes intracellulaires comme *Brucella* ou *Legionella* (Lerouge et Vanderleyden, 2002 ; Busset *et al.*, 2017). Les gènes requis pour la biosynthèse de VLCFA sont au nombre de cinq (Ardissone *et al.*, 2011). Les phénotypes des mutants *acpXL (*protéine porteuse d'acyle) et *lpxXL* (acyltransférase) ont été étudiés chez *E. meliloti* et ne sont pas capables d'effectuer une symbiose fonctionnelle avec *M. sativa*. De plus, les mutants résistent moins à divers stress abiotiques (Haag *et al.*, 2009). Seul le mutant *lpxXL* a été obtenu chez *Bradyrhizobium* sp. ORS278 et il présente un phénotype symbiotique avec problèmes de différenciation et de fixation d'azote en association avec *A. indica* et *A. evenia* (Busset *et al.*, 2017).

En dehors des VLCFA, sans doute requis pour la vie intracellulaire, l'importance des gènes liés à la biosynthèse d'hopanoïdes ou du LPS n'est donc pas forcément liée au processus symbiotique. Ces gènes contribuent sans doute à l'amélioration de la valeur sélective de la bactérie quel que soit son environnement. Cette symbiose atypique mêlant peptides NCR-like et quasi-absence de différenciation pourrait donc probablement être liée aux spécificités de

l'enveloppe de la souche *Bradyrhizobium diazoefficiens* USDA110, et plus particulièrement à la quantité d'hopanoïdes.

Perspectives d'études étendues à d'autres associations symbiotiques : Comparaisons basées sur des technologies -omiques de séquençage massif appliqué à des légumineuses d'intérêt agronomique ?

Les observations cytologiques et multi-omiques comparant les bactéroïdes de la souche *B. diazoefficiens* USDA110 entre le soja et *A. afraspera* ont mené à l'observation d'une différenciation atypique des bactéroïdes de *A. afraspera*. Cette symbiose n'a néanmoins pas été décrite dans la nature jusqu'à maintenant et est donc purement artificielle, et le processus symbiotique avec *A. afraspera* n'est pas aussi efficace que peut l'être la souche ORS285 avec le même hôte (cf. chapitre 4). Bien que les travaux effectués sur cette association n'aient pas mené à envisager un moyen d'améliorer le processus symbiotique, la transposition de ceux-ci vers des modèles d'intérêt agronomiques est relativement aisée. En effet, Oono et Denison (2010) montrent une efficacité symbiotique accrue lorsque les bactéroïdes sont différenciés en utilisant deux souches associées chacune à deux légumineuses d'intérêt agronomique où l'une induit une différenciation des bactéroïdes et l'autre non. D'un côté, la souche *Rhizobium leguminosarum* A34 s'associe au pois ou au haricot, et d'autre part *Bradyrhizobium* sp. 32H1 s'associe à l'arachide et au niébé (Oono et Denison., 2010). Le résultat de cette étude était justement à l'origine du projet visant à comparer l'efficacité symbiotique des bactéroïdes différenciés allongés ou sphériques dans le système *Bradyrhizobium-Aeschynomene* (cf. chapitre 1).

On pourrait alors imaginer effectuer des comparaisons de transcriptomes bactérien et végétal en utilisant les associations symbiotiques susmentionnées. Les génomes des hôtes végétaux sont également disponibles pour le niébé (Muñoz-Amatriaín *et al.*, 2017) et pour le haricot (Schmutz *et al.*, 2014), ainsi que pour les ancêtres diploïdes de l'arachide cultivée (Bertioli *et al.*, 2016). Enfin, le génome du pois est en voie d'être disponible (https://www.france-genomique.org/spip/spip.php?article141). L'intérêt serait à la fois de détecter les gènes impliqués dans le processus de fixation de l'azote en en prenant en compte les gènes induits dans les nodosités par rapport aux racines de chaque plante ; puis il serait intéressant de comparer autant que faire se peut les niveaux d'expression des gènes orthologues des plantes hôtes en symbiose avec la même bactérie. Du côté bactérien, les génomes des

souches A34 et 32H1 n'ont pas encore été séquencés, bien que ces souches aient été utilisées dans un certain nombre d'études datant d'avant les années 2000. Heureusement, le séquençage et l'assemblage de génomes Procaryotes n'est actuellement plus du tout un obstacle technologique et est devenu peu couteux. L'utilisation combinée des logiciels d'annotation automatique comme EugeneP-P pouvant être enrichi par des données RNA-seq (Sallet *et al.*, 2014), et de la plateforme de génomique comparative MAGE, convient parfaitement à une annotation rapide et de bonne qualité de ces génomes bactériens (Vallenet *et al.*, 2017). Dans le cadre des comparaisons transcriptomiques bactériennes, il faudrait ajouter une condition de culture bactérienne de référence en milieu minimum et en début de phase stationnaire. Ce choix de condition de référence limiterait le nombre de gènes différentiellement exprimés, dont un certain nombre étaient probablement des « faux positifs » liés à une culture en phase exponentielle en milieu riche, donc probablement beaucoup trop éloignée des conditions bactéroïdes (cf. chapitres 2, 3 et 4). Cela mènerait d'établir deux listes de gènes induits dans les bactéroïdes et d'autres spécifiques de l'hôte pouvant expliquer les différences d'efficacité symbiotique observées, de la même manière que les études présentées au cours de cette thèse.

La recherche de nouveaux gènes candidats d'intérêt pourrait également être facilitée par des expériences de Tn-seq sur ces souches bactériennes d'abord dans des cultures liquides en milieux riches et minimum en utilisant un transposon de type Mariner qui s'insère dans les sites TA du génome comme cela a été fait chez *Agrobacterium*, entre autres (Gonzalez-Mula *et al.*, accepté dans New Phytologist ; Annexe 3). En effet, cela éliminerait immédiatement les gènes létaux visibles par l'absence ou le faible nombre de reads séquencés pour ces gènes, et ainsi ne pas gaspiller de temps et d'argent à procéder à une mutagénèse impossible.

Améliorer l'efficacité symbiotique, un jalon nécessaire à l'agriculture de demain ?

e rs ec ives d'améliora ion de l'e icaci é s mbio ique des lé um ineuses

Même si optimiser la fixation de l'azote reste encore du domaine des projets non aboutis, les applications possibles et lointaines de ces travaux de thèse peuvent concerner plusieurs secteurs. Un premier exemple pourrait entrevoir les semences futures de légumineuses génétiquement améliorées pour imposer un processus de différenciation à leurs symbiontes, déplaçant au maximum l'équilibre de gain de la symbiose en faveur du partenaire végétal. Cela passerait par exemple par la transformation génétique de légumineuses devenant productrices de peptides NCR-like permettant l'amélioration de l'efficacité symbiotique de légumineuses

non productrices de NCR et agronomiquement importantes comme le soja, le haricot ou le niébé. Cela étant, il serait nécessaire d'élucider de nombreuses problématiques comme la caractérisation d'un ensemble minimal de peptides NCR augmentant de façon significative l'efficacité de la symbiose de manière robuste en prenant en compte les spécificités des autres espèces de légumineuses d'intérêt.

De surcroit, la compréhension des mécanismes d'action de ces mêmes peptides NCR laisse entrevoir une deuxième application envisageable. Il s'agirait en effet de transformer génétiquement les bactéries symbiotiques pour mimer l'effet des NCR sur les propriétés physico-chimiques des bactéroïdes, ou directement muter les gènes-cibles des NCR ou d'y ajouter des effecteurs menant directement à l'effet escompté de ces peptides sans qu'ils ne soient présents. Cela serait par exemple le cas de la protéine FtsZ dont la polymérisation est impliquée dans la formation de l'anneau-Z requis pour la division cellulaire ainsi que la chaperonne moléculaire GroEL1, qui interagissent directement avec le NCR247 (Farkas *et al.*, 2014).

Une autre piste à explorer concerne le cycle cellulaire, qui est modifié au cours de la différenciation, passant à des cycles d'endoréplication. CtrA et son régulon sont de plus en plus réprimés au cours de la différenciation spatiale des bactéroïdes s'opérant dans les nodosités indéterminées de *Medicago* (Roux *et al.*, 2014). Le même résultat a été observé lors d'un traitement au NCR247 (Penterman *et al.*, 2014). La protéine CtrA est elle-même très sujette à de nombreux niveaux de régulation (Figure 54) Il serait intéressant de moduler/mimer ce phénomène de différenciation en jouant sur les acteurs de sa régulation et possiblement sur l'efficacité du processus symbiotique. Cependant, le lien fonctionnel entre les peptides NCR et CtrA n'est pas connu et est l'enjeu du projet ANR SymbiontCellCyc (2017-2020) entre notre équipe et celle d'Emanuele Biondi.

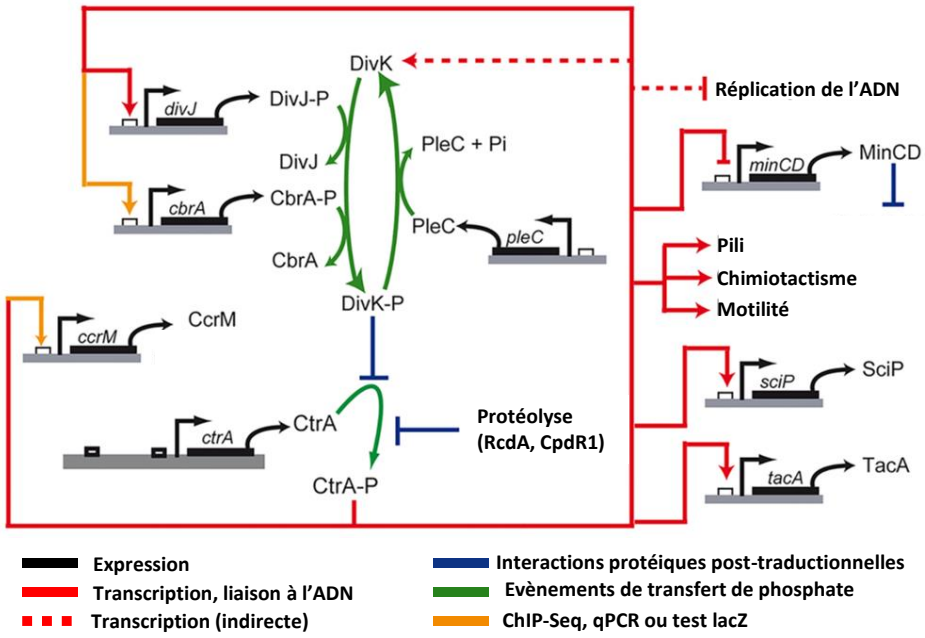


Figure 54: Boucles de régulation médiées par CtrA et démontrées via diverses méthodes (d'après Pini *et al.*, 2015).

e rans er de la i a ion d'azo e au non-légumineuses

L'avantage sélectif apporté par le processus symbiotique entre les rhizobia et les légumineuses n'est plus à prouver, tant au niveau écologique qu'agronomique. Les potentialités du transfert artificiel de la fixation d'azote résultant de l'étude de cette symbiose sont énormes et multiples, et donc des moteurs actifs de la recherche scientifique internationale. De nombreux financements affluent dans le but d'améliorer la contribution de cette fixation symbiotique de l'azote, provenant notamment d'agences de recherche comme l'ERC (European Reasearch Council) et de fondations pour la recherche (Fondation Bill and Melinda Gates). Au travers de l'observation des processus symbiotiques et de la compréhension de leurs mécanismes, nous contribuons à la mise en place de technologies successives de plus en plus abouties pouvant mener à un transfert fonctionnel et efficace de la capacité de fixation d'azote aux non-légumineuses (Mus *et al.*, 2016).

Cet intérêt est d'abord focalisé sur les Poacées telles que le blé, le riz ou le maïs servant comme matière première à l'alimentation de base des humains depuis des millénaires et ce, avant même l'invention de l'agriculture dans le cas du pain comme semblent le montrer de récentes fouilles archéologiques (Arranz-Otaegui *et al.*, 2018). En premier lieu, beaucoup de recherches ont été effectuées dans le but de transférer la capacité à mettre en place des nodosités fixatrices d'azote à des non-légumineuses (Rogers & Oldroyd, 2014). En effet, au cours des 30

dernières années, la communauté scientifique est parvenue à décortiquer de plus en plus précisément le dialogue moléculaire complexe des étapes précoces du processus de nodulation (Oldroyd, 2013).

La mise en évidence de cette voie de signalisation a été ensuite à la base de découvertes spectaculaires qui ont suscité un regain d'optimisme auprès des partisans du transfert de la nodulation aux non-Légumineuses, qui poursuivent actuellement cet ambitieux objectif à long terme. Ces avancées incluent la découverte de l'implication d'éléments génétiques de la voie de signalisation des facteurs Nod dans d'autres processus endosymbiotiques comme la symbiose actinorhizienne ou la mycorhization arbusculaire comme précédemment développé dans l'introduction (Markmann & Parniske, 2009 ; Oldroyd, 2013). Comme la mycorhization concerne plus de 85% des plantes terrestres, dont les Poacées, cela signifie que certains éléments de la voie de signalisation des facteurs Nod sont déjà présents et fonctionnels chez des non-Légumineuses. La stratégie d'ingénierie génétique de la nodulation dans les plantes cultivées pourrait donc en être simplifiée en tirant parti des voies de signalisation préexistantes. La diversité des processus d'organogénèse des légumineuses est également à prendre en compte. En effet, le mécanisme de formation des nodosités indéterminées des IRLC, bien qu'étant très étudié, est extrêmement complexe notamment au travers des exemples concernant les différents rôles des peptides NCR. Bien que cette mécanique complexe mène surement à une amélioration des gains en azote pour la plante tout en contrôlant la population des bactéroïdes, il serait intéressant d'envisager une diversification des modèles d'étude qui permettrait d'identifier un système minimal de gènes pouvant induire la formation de nodosités. L'amélioration de l'efficacité de la fixation de l'azote se ferait alors dans un second temps par ingénierie génétique, mimant et accélérant les processus évolutifs observés.

Des pistes alternatives pour le transfert de la fixation d'azote aux non-légumineuses ?

D'autres pistes de recherches sont également l'objet d'investigations pour transférer la fixation d'azote aux autres plantes. La première serait d'utiliser des bactéries endophytes diazotrophes. Une autre piste serait d'utiliser les connaissances sur la biochimie de la fixation d'azote pour transférer un complexe fonctionnel de la nitrogénase dans les organites des végétaux, qui seraient alors appelés « nitroplastes ».

U ilisa ion d'endo h es diazo ro he s

Parmi les voies explorées actuellement, l'utilisation des bactéries endophytes fixatrices d'azote (Geddes *et al.*, 2015) semble *a priori* la solution la plus simple. Cette solution nécessite dans les faits de trouver le meilleur compromis entre l'efficacité de colonisation endophytique et le taux de fixation d'azote. En effet, des souches d'endophytes fixateurs d'azote ont été isolées à partir de nombreuses plantes cultivées, et le criblage de souches présentant une efficacité de fixation élevée devrait mener à l'isolation de nouvelles souches d'intérêt. Un exemple déjà mentionné est celui de la canne à sucre colonisée par *Gluconacetobacter diazotrophicus* dont le gain de masse en azote peut dépasser les 50% grâce à la fixation d'azote des endophytes (Eskin *et al.*, 2014).

Cependant, ce cas est loin d'être une généralité, et souvent l'azote n'est d'ailleurs transféré à la plante qu'en quantité très basses même si des effets bénéfiques sont observés (Santi *et al.*, 2013). Une solution serait de transférer la capacité de fixation d'azote à une souche colonisant très efficacement les tissus végétaux. Cette approche a été réalisée avec succès en laboratoire, en transférant les gènes de fixation d'azote de la souche diazotrophe *Pseudomonas stutzeri* vers l'endophyte non fixateur *Pseudomonas protegens* Pf-5, donnant la souche Pf-5 X940 (Setten *et al.*, 2013). La souche présentait une forte activité nitrogénase dans les racines des plantes hôtes, libérant ainsi beaucoup d'azote assimilable par la plante. L'effet sur la croissance des plantes hôtes est très positif, avec une forte colonisation des racines, mais une persistance dans le sol décevante (Fox *et al.*, 2016). Enfin, une découverte récente a mis en lumière l'existence d'une association mutualiste inédite entre des bactéries diazotrophes et une espèce de maïs native de la région Sierra Mixe au Mexique vivant sur des sols déplétés en azote. Cette association naturelle se caractérise par le développement de racines aériennes sécrétant un mucilage riche en sucres (Van Deynze *et al.*, 2018). La fixation d'azote de ces bactéries semble être responsable d'entre 30% et 80% de la nutrition azotée de cette espèce de maïs, donnant ainsi une nouvelle piste visant à obtenir des céréales ne nécessitant pas d'intrants azotés pour se développer.

La création de « nitroplastes »

Une dernière stratégie consiste à introduire directement dans la plante les gènes codant la nitrogénase (Curatti & Rubio, 2014). Du fait de sa structure complexe précédemment décrite dans l'introduction, l'assemblage et le fonctionnement de la nitrogénase pose la première

difficulté de nécessiter la participation de nombreux gènes (e.g., 18 chez *Klebsiella oxytoca*). Parmi eux, il y a les gènes codant les sous-unités structurales de la nitrogénase, mais aussi les enzymes impliquées dans la biogenèse des clusters métalliques, du FeMo-Co et dans leur insertion dans NifDK. Une autre limite est le besoin d'être en situation micro-oxique pour éviter l'oxydation des cofacteurs métalliques de l'enzyme. La toxicité cellulaire de l'ammonium produit requiert une assimilation rapide de ce dernier, et peut nécessiter une ou plusieurs voies métaboliques le prenant en charge au niveau de son lieu de production. Finalement, la nécessité d'avoir beaucoup d'ATP et de pouvoir réducteur pour que la cinétique de la réaction soit suffisamment rapide impose l'expression de la nitrogénase dans des organites produisant de l'ATP en masse, à savoir les mitochondries et les chloroplastes. Ce type d'application biotechnologique fait donc face à un certain nombre de verrous technologiques complexes à résoudre.

L'utilisation des organites Procaryotes autoriserait l'utilisation directe de constructions géniques basées sur des promoteurs bactériens et constituées en opérons. Les chloroplastes présentent l'avantage d'être aisément transformables génétiquement. Néanmoins, l'activité photosynthétique génératrice de dioxygène des chloroplastes semble difficilement compatible avec la non-dénaturation d'une activité nitrogénase. Une solution serait d'avoir de la fixation d'azote seulement la nuit, mais avec un rendement moindre vu qu'il n'y aurait pas de régénération d'ATP et de pouvoir réducteur provenant de la phase photochimique de la photosynthèse. La stratégie envisagée pour transformer des mitochondries en nitroplastes est plutôt une transformation de la cellule végétale qui produirait les protéines du complexe nitrogénase qui seraient ensuite adressées aux mitochondries.

Des expériences ont été effectuées en utilisant des levures transformées produisant et adressant NifH, et montrant l'existence d'un recrutement indépendant de clusters Fer-Soufre mitochondriaux ne nécessitant alors pas les protéines NifS et NifU (Lopez-Torrejon *et al.*, 2016). Cette approche a aussi été récemment testée en transformant du tabac avec les gènes *nif* de *Klebsiella oxytoca* (Allen *et al.*, 2017). Il existe pourtant d'autres complexes nitrogénase fonctionnels comportant moins de gènes que celui de *K. oxytoca*, comme par exemple *Paenibacillus* sp. WLY78 qui n'en possède que 9 (*nifBHDKENXV* et *hesA*) et qui pourraient être utilisés en substituts pour simplifier le processus de transformation génétique et d'assemblage du complexe nitrogénase (Wang *et al.*, 2013).

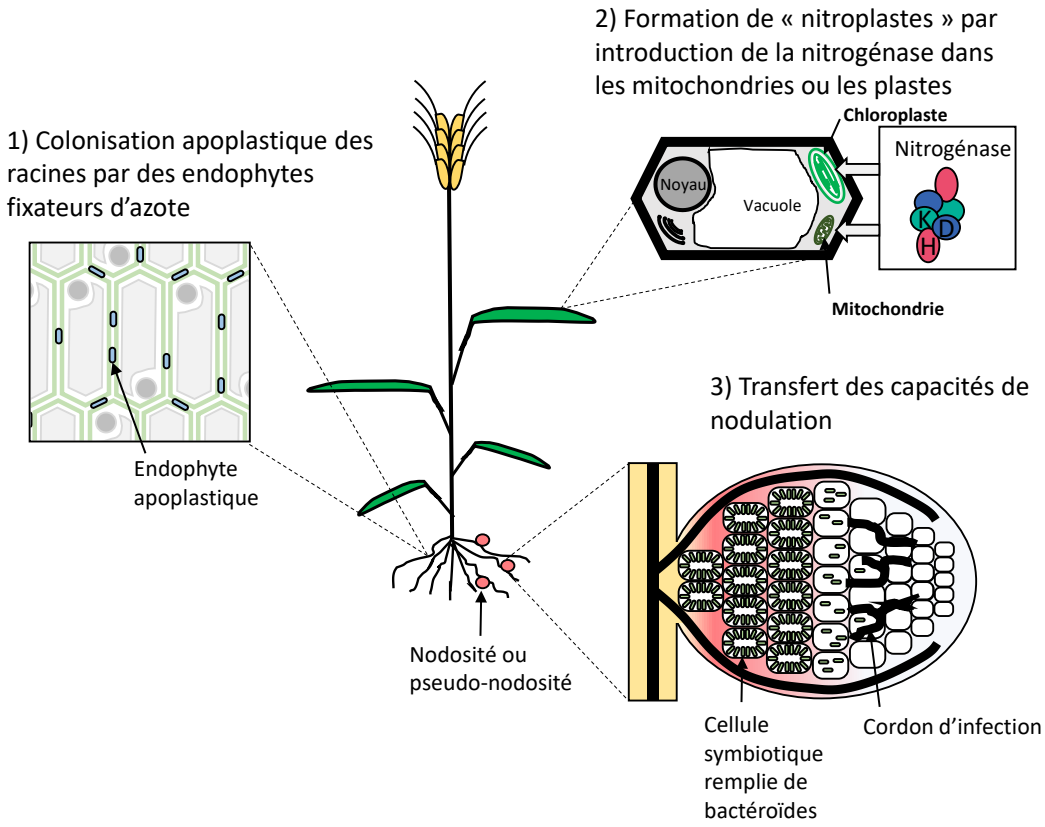


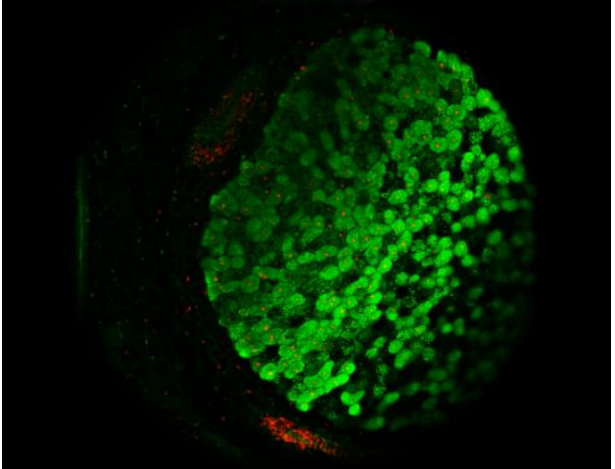
Figure 55: Stratégies actuelles visant à transférer la fixation biologique de l'azote chez des plantes de grande culture autres hors-légumineuses. 1) Des souches endophytes diazotrophes colonisant les racines des plantes pourront être testées pour leur contribution à la nutrition azotée de l'hôte. 2) La création de « nitroplastes » reconstituant la voie métabolique de la fixation biologique de l'azote par introduction de gènes codant le complexe nitrogénase dans les plastes et/ou les mitochondries. 3) L'introduction des gènes contrôlant la formation et le fonctionnement de nodosités pourrait mener à la mise en place de nodosités fixatrices d'azote sur les systèmes racinaires de plantes de grande culture (adaptée de Alunni et Mergaert, 2017).

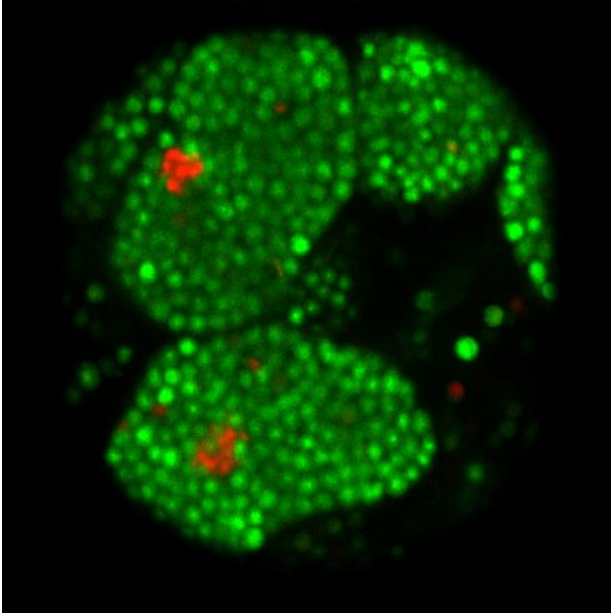
Conclusion

Les travaux effectués durant cette thèse ont permis d'approfondir nos connaissances concernant l'adaptation physiologique des bactéries du genre *Bradyrhizobium* au processus de différenciation imposé par les plantes hôtes du genre *Aeschynomene* au travers d'approches globales sans *a priori*. Le système expérimental initial disposait d'un certain attrait avec un gradient de différenciation allant des bactéroïdes du soja à ceux de *A. indica* permettant en théorie de définir les signatures moléculaires du processus de différenciation le long de ce gradient. Dans la réalité, nous nous sommes heurtés à divers problèmes. Tout d'abord l'utilisation de deux espèces bactériennes différentes rend très difficile l'intégration complète de l'ensemble des jeux de données produits et analysés au cours de cette thèse. Au sein même de ce gradient, les valeurs intermédiaires prises par les bactéroïdes des souches de ORS285 ou

USDA110 de *A. afraspera* semblent discutables. Elles nécessitent sans doute d'être redéfinies avant de trop s'avancer pour parler de processus de différenciation similaire à celui des IRLC chez *A. afraspera*. Cela vaut donc également pour les nouvelles pistes sur la corrélation possible entre le processus de différenciation des bactéroïdes et l'efficacité de la symbiose qui ont été suivies. Toujours est-il que bien malgré le doute quant à l'existence d'une différenciation des bactéroïdes d'*A. afraspera*, celle des symbiontes des *Aeschynomene* du groupe III ne fait aucun doute et présente une efficacité symbiotique accrue. Il conviendrait de pousser l'analyse à des stades phénologiques plus avancés pour être conclusif.

Au cours de ce travail, j'ai effectué des analyses descriptives des jeux de données en comparant les génotypes bactériens sauvages, mais également fonctionnelles en observant l'effett de la mutation de bclA sur le transcriptome bactérien en vie libre et symbiotique. J'ai également réalisé des études fonctionnelles passant par la création de mutant bactériens et le phénotypage d'un grand nombre d'interactions spécifiques sur les associations *Bradyrhizobium-Aeschynomene*, mais également sur le modèle commun *Ensifer-Medicago*. Ces travaux nous ont permis de mieux appréhender la complexité et la robustesse de cette adaptation physiologique des bactéries à la condition symbiotique au travers de l'absence de phénotype observé des nombreux mutants générés. Le décryptage de cette robustesse passera par la génération de mutants multiples permettant de cartographier ces complexes réseaux de gènes responsables du processus de différenciation terminale des bactéroïdes. Ce dernier pourra devenir le sixième processus prometteur à intégrer après la reconnaissance des partenaires, l'initiation de l'infection, l'organogenèse nodulaire, intracellulaire des bactéries et leur intégration métabolique au sein de l'hôte légumineuse ou non.





Références bibliographiques

Références bibliographiques

Abeysekera, R.M., Newcomb, W., Silvester, W.B., and Torrey, J.G. (1990) A freeze-fracture electron microscopic study of *Frankia* in root nodules of *Alnus incana* grown at three oxygen tensions. *Can J Microbiol* **36**: 97–108.

Alazard, D. (1990) Nitrogen fixation in pure culture by rhizobia isolated from stem nodules of tropical *Aeschynomene* species. *FEMS Microbiol Lett* **68**: 177–182.

Alazard, D. (1985) Stem and root nodulation in *Aeschynomene* spp. *Appl Environ Microbiol* **50**: 732–734.

Allen, R.S., Tilbrook, K., Warden, A.C., Campbell, P.C., Rolland, V., Singh, S.P., and Wood, C.C. (2017) Expression of 16 Nitrogenase Proteins within the Plant Mitochondrial Matrix. *Front Plant Sci* **8**: 1–14.

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.

Alunni, B. and 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.

Amann, M., Dick, D., Bertil, F., Hänninen, O., Hurley, F., Krzyzanowski, M., *et al.* (2008) Health risks of ozone from long-range transboundary air pollution. *World Heal Organ* 111 p.

Ampe, F., Kiss, E., Sabourdy, F., and Batut, J. (2003) Transcriptome analysis of *Sinorhizobium meliloti* during symbiosis. *Genome Biol* **4**: R15.

Ané, J.M., Kiss, G.B., Riely, B.K., Penmetsa, R.V., Oldroyd, G.E.D., Ayax, C., *et al.* (2004) *Medicago truncatula* DMI1 required for bacterial and fungal symbioses in legumes. *Science* **303**: 1364–1367.

Appleby, C.A., Tjepkema, J.D., and Trinick, M.J. (1983) Hemoglobin in a nonleguminous plant, *Parasponia*: possible genetic origin and function in nitrogen fixation. *Science* **220**: 951–953.

Ardissone, S., Kobayashi, H., Kambara, K., Rummel, C., Noel, K.D., Walker, G.C., *et al.* (2011) Role of BacA in lipopolysaccharide synthesis, peptide transport, and nodulation by Rhizobium sp. strain NGR234. *J Bacteriol* **193**: 2218–2228.

Arranz-Otaegui, A., Gonzalez Carretero, L., Ramsey, M.N., Fuller, D.Q., and Richter, T. (2018) Archaeobotanical evidence reveals the origins of bread 14,400 years ago in northeastern Jordan. *Proc Natl Acad Sci U S A* **115**: 7925–7930.

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.

Baetz, U. and Martinoia, E. (2014) Root exudates: The hidden part of plant defense. *Trends Plant Sci* **19**: 90–98.

Barnett, M.J., Toman, C.J., Fisher, R.F., and Long, S.R. (2004) A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. *Proc Natl Acad Sci U S A* **101**: 16636–16641.

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.

Becker, A., Berges, H., Krol, E., Bruand, C., Ruberg, S., Capela, D., *et al.* (2004) Global changes in gene expression in *Sinorhizobium meliloti* 1021 under microoxic and symbiotic conditions. *Mol Plant Microbe Interact* **17**: 292–303.

Beijerinck, M.W. (1888) Die Bacterien der Papilionaceenknöllchen. *Bot Zeitung* **46**: 725–804.

Belin, B.J., Busset, N., Giraud, E., Molinaro, A., Silipo, A., and Newman, D.K. (2018) Hopanoid lipids: From membranes to plant-bacteria interactions. *Nat Rev Microbiol* **16**: 304–315.

Berman-Frank, I., Lundgren, P., Chen, Y.B., Küpper, H., Kolber, Z., Bergman, B., and Falkowski, P. (2001) Segregation of nitrogen fixation and oxygenic photosynthesis in the marine cyanobacterium *Trichodesmium*. *Science* **294**: 1534–1537.

Bernhard, A. (2010) The Nitrogen cycle: processes, players, and human impact. *Nat Educ Knowl* **3(10)**: 25.

Berrabah, F., Ratet, P., and Gourion, B. (2015) Multiple steps control immunity during the intracellular accommodation of rhizobia. *J Exp Bot* **66**: 1977–1985.

Berry, A.M., Harriott, O.T., Moreau, R.A., Osman, S.F., Benson, D.R., and Jones, A.D. (1993) Hopanoid lipids compose the *Frankia* vesicle envelope, presumptive barrier of oxygen diffusion to nitrogenase. *Proc Natl Acad Sci U S A* **90**: 6091–6094.

Bertioli, D.J., Cannon, S.B., Froenicke, L., Huang, G., Farmer, A.D., Cannon, E.K.S., *et al.* (2016) The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nat Genet* **48**: 438–446.

Bhattacharyya, P.N. and Jha, D.K. (2012) Plant growth-promoting rhizobacteria (PGPR): Emergence in agriculture. *World J Microbiol Biotechnol* **28**: 1327–1350.

Biondi, E.G., Reisinger, S.J., Skerker, J.M., Arif, M., Perchuk, B.S., Ryan, K.R., and Laub, M.T. (2006) Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature* **444**: 899–904.

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.

Bonaldi, K., Gourion, B., Fardoux, J., Hannibal, L., Cartieaux, F., Boursot, M., *et al.* (2010) Large-scale transposon mutagenesis of photosynthetic *Bradyrhizobium* sp. strain ORS278 reveals new genetic loci putatively important for Nod-independent symbiosis with *Aeschynomene indica*. *Mol Plant-Microbe Interact* **23**: 760–770.

Borucki, W.J. and Chameides, W.L. (1984) Lightning: Estimates of the rates of energy dissipation and nitrogen fixation. *Rev Geophys* **22**: 363–372.

Boulaine, J. (1995a) Quatre siècles de fertilisation : première partie. *Etude Gest des sols* **2**: 201–208.

Boulaine, J. (1995b) Quatre siècles de fertilisation: seconde partie. *Etude Gest des sols* **2**: 219–226.

Boyd, E.S., Anbar, A.D., Miller, S., Hamilton, T.L., Lavin, M., and Peters, J.W. (2011) A late methanogen origin for molybdenum-dependent nitrogenase. *Geobiology* **9**: 221–232.

Boyd, E.S. and Peters, J.W. (2013) New insights into the evolutionary history of biological nitrogen fixation. *Front Microbiol* **4**: 1–12.

Broughton, W.J., Jabbouri, S., and Perret, X. (2000) Keys to symbiotic harmony. *J Bacteriol* **182**: 5641–5652.

Bruneau, A., Doyle, J.J., Herendeen, P., Hughes, C.E., Kenicer, G., Lewis, G., *et al.* (2013) Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon* **62**: 217–248.

Busset, N., De Felice, A., Chaintreuil, C., Gully, D., Fardoux, J., Romdhane, S., *et al.* (2016) The LPS O-antigen in photosynthetic *Bradyrhizobium* strains is dispensable for the establishment of a successful symbiosis with *Aeschynomene* legumes. *PLoS One* **11**: 1–14.

Busset, N., Lorenzo, F. Di, Palmigiano, A., Sturiale, L., Gressent, F., Fardoux, J., *et al.* (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. *Front Microbiol* **8**: 1–13.

Cao, Y., Halane, M.K., Gassmann, W., and Stacey, G. (2017) The Role of Plant Innate Immunity in the Legume-Rhizobium Symbiosis. *Annu Rev Plant Biol* **68**: 535–561.

Capela, D., Carrere, S., and Batut, J. (2005) Transcriptome-based identification of the *Sinorhizobium meliloti* NodD1 regulon. *Appl Environ Microbiol* **71**: 4910–4913.

Capela, D., Filipe, C., Bobik, C., Batut, J., and Bruand, C. (2006) *Sinorhizobium meliloti* differentiation during symbiosis with alfalfa: A transcriptomic dissection. *Mol Plant-Microbe Interact* **19**: 363–372.

Capoen, W., Sun, J., Wysham, D., Otegui, M.S., Venkateshwaran, M., Hirsch, S., *et al.* (2011) Nuclear membranes control symbiotic calcium signaling of legumes. *Proc Natl Acad Sci U S A* **108**: 14348–14353.

Carlier, A., Chevrot, R., Dessaux, Y., and Faure, D. (2004) The assimilation of gamma-butyrolactone in *Agrobacterium tumefaciens* C58 interferes with the accumulation of the N-acyl-homoserine lactone signal. *Mol Plant Microbe Interact* **17**: 951–7.

Carlson, R.W., Hanley, B., Rolfe, B.G., and Djordjevic, M.A. (1986) A Structural comparison of the acidic extracellular polysaccharides from *Rhizobium trifolii* mutants affected in root hair infection. *Plant Physiol* **80**: 134–7.

Carro, L., Pujic, P., Alloisio, N., Fournier, P., Boubakri, H., Hay, A.E., *et al.* (2015) *Alnus* peptides modify membrane porosity and induce the release of nitrogen-rich metabolites from nitrogen-fixing *Frankia*. *ISME J* **9**: 1723–1733.

Cebolla, A., Vinardell, J.M., Kiss, E., Oláh, B., Roudier, F., Kondorosi, A., and Kondorosi, E. (1999) The mitotic inhibitor *ccs52* is required for endoreduplication and ploidy-dependent cell enlargement in plants. *EMBO J* **18**: 4476–4484.

Chai, Y., Tsai, C.S., Cho, H., and Winans, S.C. (2007) Reconstitution of the biochemical activities of the AttJ repressor and the AttK, AttL, and AttM catabolic enzymes of *Agrobacterium tumefaciens*. *J Bacteriol* **189**: 3674–9.

Chaintreuil, C., Arrighi, J.F., Giraud, E., Miché, L., Moulin, L., Dreyfus, B., *et al.* (2013) Evolution of symbiosis in the legume genus *Aeschynomene*. *New Phytol* **200**: 1247–1259.

Chaintreuil, C., Gully, D., Hervouet, C., Tittabutr, P., Randriambanona, H., Brown, S.C., *et al.* (2016) The evolutionary dynamics of ancient and recent polyploidy in the African semiaquatic species of the legume genus *Aeschynomene*. *New Phytol* **211**: 1077–1091.

Charpentier, M., Sun, J., Martins, T.V., Radhakrishnan, G. V, Findlay, K., Soumpourou, E., *et al.* (2016) Nuclear-localized cyclic nucleotide–gated channels mediate symbiotic calcium oscillations. *Science* **352**: 1102 LP-1105.

Coba de la Peña, T., Fedorova, E., Pueyo, J.J., and Lucas, M.M. (2018) The Symbiosome: Legume and Rhizobia Co-evolution toward a Nitrogen-Fixing Organelle? *Front Plant Sci* **8**: 1–26.

Cole, B.J., Feltcher, M.E., Waters, R.J., Wetmore, K.M., Mucyn, T.S., Ryan, E.M., *et al.* (2017) Genome-wide identification of bacterial plant colonization genes. *PLoS Biol* **15**: 1–24.

Cren, M., Kondorosi, A., and Kondorosi, E. (1995) NoIR controls expression of the *Rhizobium meliloti* nodulation genes involved in the core Nod factor synthesis. *Mol Microbiol* **15**: 733–747.

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.

Curatti, L. and Rubio, L.M. (2014) Challenges to develop nitrogen-fixing cereals by direct nif-gene transfer. *Plant Sci* **225**: 130–7.

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., *et al.* (2015) Convergent evolution of rndosymbiont differentiation in Dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by nodule-specific cysteine-rich peptides. *Plant Physiol* **169**: 1254–1265.

Davidson, A.L., Dassa, E., Orelle, C., and Chen, J. (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* **72**: 317 LP-364.

Davidson, E.A. (2009) The contribution of manure and fertilizer nitrogen to atmospheric nitrous oxide since 1860. *Nat Geosci* **2**: 659–662.

Deakin, W.J. and Broughton, W.J. (2009) Symbiotic use of pathogenic strategies: Rhizobial protein secretion systems. *Nat Rev Microbiol* **7**: 312–320.

Delgado, M.J., Bedmar, E.J., and Downie, J.A. (1998) Genes involved in the formation and assembly of rhizobial cytochromes and their role in symbiotic nitrogen fixation., pp. 191–231.

Demoule, J.P. (2010) La révolution néolithique dans le monde. *CNRS Ed Paris, Fr* 488 p.

Deutschbauer, A., Price, M.N., Wetmore, K.M., Shao, W., Baumohl, J.K., Xu, Z., *et al.* (2011) Evidence-based annotation of gene function in *Shewanella oneidensis MR-1* using genome-wide fitness profiling across 121 conditions. *PLOS Genet* **7**: e1002385.

Van Deynze, A., Zamora, P., Delaux, P.-M., Heitmann, C., Jayaraman, D., Rajasekar, S., *et al.* (2018) Nitrogen fixation in a landrace of maize is supported by a mucilage-associated diazotrophic microbiota. *PLOS Biol* **16**: e2006352.

diCenzo, G., Milunovic, B., Cheng, J., and Finan, T.M. (2013) The tRNAarg 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. *J Bacteriol* **195**: 202–212.

diCenzo, G.C., Zamani, M., Checcucci, A., Fondi, M., Griffitts, J.S., Finan, T.M., and Mengoni, A. (2018) Multi-disciplinary approaches for studying rhizobium – legume symbioses. *Peer J* cjm-2018-0377.

Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., *et al.* (2013) A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform* **14**: 671–83.

Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J. (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* **2**: 414–424.

Domenech, P., Kobayashi, H., LeVier, K., Walker, G.C., and Barry, C.E. (2009) BacA, an ABC transporter involved in maintenance of chronic murine infections with mycobacterium tuberculosis. *J Bacteriol* **191**: 477–485.

Dordas, C., Rivoal, J., and Hill, R.D. (2003) Plant haemoglobins, nitric oxide and hypoxic stress. *Ann Bot* **91**: 173–178.

Dupont, L., Alloing, G., Pierre, O., El, S., Hopkins, J., Hrouart, D., and Frendo, P. (2012) The Legume root nodule: From symbiotic nitrogen fixation to senescence. In, *Senescence*. InTech.

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.

Eskin, N., Vessey, K., and Tian, L. (2014) Research progress and perspectives of nitrogen fixing bacterium, *Gluconacetobacter diazotrophicus* , in monocot plants. *Int J Agron* **2014**: 1–13.

Fabre, S., Gully, D., Poitout, A., Patrel, D., Arrighi, J.-F., Giraud, E., *et al.* (2015) The Nod factor-independent nodulation in *Aeschynomene evenia* required the common plant-microbe symbiotic "toolkit." *Plant Physiol* **169**: pp.01134.2015.

Fagan, R.P. and Fairweather, N.F. (2014) Biogenesis and functions of bacterial S-layers. *Nat Rev Microbiol* **12**: 211–22.

Farkas, A., Maroti, G., Durg, H., Gyorgypal, Z., Lima, R.M., Medzihradszky, K.F., *et al.* (2014) *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proc Natl Acad Sci U S A* **111**: 5183–5188.

Fay, P. (1992) Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev* **56**: 340–73.

Fields, S. (2004) Global nitrogen: cycling out of control. *Environ Health Perspect* **112**: 556–563.

Fiore, C.L., Jarett, J.K., Olson, N.D., and Lesser, M.P. (2010) Nitrogen fixation and nitrogen transformations in marine symbioses. *Trends Microbiol* **18**: 455–463.

Fox, A.R., Soto, G., Valverde, C., Russo, D., Lagares, A., Zorreguieta, Á., *et al.* (2016) Major cereal crops benefit from biological nitrogen fixation when inoculated with the nitrogen-fixing bacterium Pseudomonas protegens Pf-5 X940. *Environ Microbiol* **18**: 3522–3534.

Futahashi, R., Tanaka, K., Tanahashi, M., Nikoh, N., Kikuchi, Y., Lee, B.L., and Fukatsu, T. (2013) Gene expression in gut symbiotic organ of stinkbug affected by extracellular bacterial symbiont. *PLoS One* **8**: e64557.

Galibert, F., Finan, T.M., Long, S.R., Pühler, A., Abola, P., Ampe, F., *et al.* (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**: 668–672.

Geddes, B.A., Ryu, M.-H., Mus, F., Garcia Costas, A., Peters, J.W., Voigt, C.A., and Poole, P. (2015) Use of plant colonizing bacteria as chassis for transfer of N_2-fixation to cereals.

Curr Opin Biotechnol **32**: 216–222.

Gherbi, H., Duhoux, E., Franche, C., Pawlowski, K., Nassar, A., Berry, A.M., and Bogusz, D. (1997) Cloning of a full-length symbiotic hemoglobin cDNA and in situ localization of the corresponding mRNA in *Casuarina glauca* root nodule. *Physiol Plant* **99**: 608–616.

Ghodhbane-Gtari, F., Hezbri, K., Ktari, A., Sbissi, I., Beauchemin, N., Gtari, M., and Tisa, L.S. (2014) Contrasted reactivity to oxygen tensions in *Frankia* sp. strain CcI3 throughout nitrogen fixation and assimilation. *Biomed Res Int* **2014**.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., *et al.* (2002) Functional profiling of the Saccharomyces cerevisiae genome. *Nature* **418**: 387–91.

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.-C., *et al.* (2007) Legumes symbioses: absence of nod genes in photosynthetic bradyrhizobia. *Science* **316**: 1307–1312.

Giraud, E., Xu, L., Chaintreuil, C., Gargani, D., Gully, D., and Sadowsky, M.J. (2013) Photosynthetic *Bradyrhizobium* sp. strain ORS285 is capable of forming nitrogen-fixing root nodules on soybeans (Glycine max). *Appl Environ Microbiol* **79**: 2459–2462.

Glazebrook, J., Ichige, A., and Walker, G.C. (1993) A *Rhizobium meliloti* homolog of the *Escherichia coli* peptide-antibiotic transport protein SbmA is essential for bacteroid development. *Genes Dev* **7**: 1485–1497.

Goldberg, I., Nadler, V., and Hochman, A. (1987) Mechanism of nitrogenase switch-off by oxygen. *J Bacteriol* **169**: 874–879.

Golden, J.W. and Yoon, H.S. (2003) Heterocyst development in *Anabaena*. *Curr Opin Microbiol* **6**: 557–563.

González-Sama, A., de la Peña, T.C., Kevei, Z., Mergaert, P., Lucas, M.M., de Felipe, M.R., *et al.* (2006) Nuclear DNA endoreduplication and expression of the mitotic inhibitor Ccs52 associated to determinate and lupinoid nodule organogenesis. *Mol Plant Microbe Interact* **19**: 173–80.

Gould, S.B., Waller, R.F., and McFadden, G.I. (2008) Plastid Evolution. *Annu Rev Plant Biol* **59**: 491–517.

Gruber, N. and Galloway, J.N. (2008) An Earth-system perspective of the global nitrogen cycle. *Nature* **451**: 293–296.

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.

Gully, D., Gargani, D., Bonaldi, K., Grangeteau, C., Chaintreuil, C., Fardoux, J.J., *et al.* (2016) A Peptidoglycan-remodeling enzyme is critical for bacteroid differentiation in *Bradyrhizobium* spp. during legume symbiosis. *Mol Plant-Microbe Interact* **29**: 447–457.

Gully, D., Teulet, A., Busset, N., Nouwen, N., Fardoux, J., Rouy, Z., *et al.* (2017) Complete genome sequence of *Bradyrhizobium* sp. ORS285, a photosynthetic strain able to establish Nod factor-eependent or Nod factor-independent symbiosis with *Aeschynomene* legumes. *Genome Announc* **5**: e00421-17.

Gyaneshwar, P., James, E.K., Mathan, N., Reddy, P.M., Reinhold-Hurek, B., and Ladha, J.K. (2001) Endophytic colonization of rice by a diazotrophic strain of *Serratia marcescens*. *J Bacteriol* **183**: 2634–2645.

Haag, A.F., Arnold, M.F.F., Myka, K.K., Kerscher, B., Dall'Angelo, S., Zanda, M., *et al.* (2013)

Molecular insights into bacteroid development during Rhizobium-legume symbiosis. *FEMS Microbiol Rev* **37**: 364–83.

Haag, A.F., Baloban, M., Sani, M., Kerscher, B., Pierre, O., Farkas, A., *et al.* (2011) Protection of *Sinorhizobium* against host cysteine-rich antimicrobial peptides is critical for symbiosis. *PLoS Biol* **9**: e1001169.

Haag, A.F., Wehmeier, S., Beck, S., Marlow, V.L., Fletcher, V., James, E.K., and Ferguson, G.P. (2009) The *Sinorhizobium meliloti* LpxXL and AcpXL proteins play important roles in bacteroid development within alfalfa. *J Bacteriol* **191**: 4681–4686.

Hakoyama, T., Niimi, K., Watanabe, H., Tabata, R., Matsubara, J., Sato, S., *et al.* (2009) Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. *Nature* **462**: 514–517.

Haskett, T.L., Terpolilli, J.J., Bekuma, A., O'Hara, G.W., Sullivan, J.T., Wang, P., *et al.* (2016) Assembly and transfer of tripartite integrative and conjugative genetic elements. *Proc Natl Acad Sci U S A* **113**: 12268–12273.

Hauser, F., Pessi, G., Friberg, M., Weber, C., Rusca, N., Lindemann, A., *et al.* (2007) Dissection of the *Bradyrhizobium japonicum* NifA+σ54 regulon, and identification of a ferredoxin gene (*fdxN*) for symbiotic nitrogen fixation. *Mol Genet Genomics* **278**: 255–271.

Hempel, J., Zehner, S., Göttfert, M., and Patschkowski, T. (2009) Analysis of the secretome of the soybean symbiont *Bradyrhizobium japonicum*. *J Biotechnol* **140**: 51–58.

Horrigan, L., Lawrence, R.S., and Walker, P. (2002) How sustainable agriculture can address the environmental and human health harms of industrial agriculture. *Environ Health Perspect* **110**: 445–456.

Horváth, B., Domonkos, Á., Kereszt, A., Szűcs, A., Ábrahám, E., Ayaydin, F., *et al.* (2015) Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the *Medicago truncatula dnf7* mutant. *Proc Natl Acad Sci U S A* **112**: 15232–15237.

Hottes, A.K., Shapiro, L., and McAdams, H.H. (2005) DnaA coordinates replication initiation and cell cycle transcription in *Caulobacter crescentus*. *Mol Microbiol* **58**: 1340–1353.

Huang, C.-T., Liu, C.-T., Chen, S.-J., and Kao, W.-Y. (2016) Phylogenetic identification, phenotypic variations, and symbiotic characteristics of the peculiar rhizobium, strain CzR2, isolated from *Crotalaria zanzibarica* in Taiwan. *Microbes Environ* **31**: 410–417.

Hur, G.H., Vickery, C.R., and Burkart, M.D. (2012) Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat Prod Rep* **29**: 1074–1098.

Ivanov, S., Fedorova, E.E., Limpens, E., De Mita, S., Genre, A., Bonfante, P., and Bisseling, T. (2012) Rhizobium-legume symbiosis shares an exocytotic pathway required for arbuscule formation. *Proc Natl Acad Sci U S A* **109**: 8316–8321.

Janson, S., Wouters, J., Bergman, B., and Carpenter, E.J. (1999) Host specificity in the Richelia-diatom symbiosis revealed by hetR gene sequence analysis. *Environ Microbiol* **1**: 431–438.

Jensen, E.S. and Sørensen, L.H. (1987) Survival of *Rhizobium leguminosarum* in soil after addition as inoculant. *FEMS Microbiol Lett* **45**: 221–226.

Jiménez-Guerrero, I., Acosta-Jurado, S., del Cerro, P., Navarro-Gómez, P., López-Baena, F.J., Ollero, F.J., *et al.* (2018) Transcriptomic studies of the effect of nod gene-inducing molecules in rhizobia: Different weapons, one purpose. *Genes (Basel)* **9**: 1–27.

Jones, K.M., Kobayashi, H., Davies, B.W., Taga, M.E., and Walker, G.C. (2007) How rhizobial symbionts invade plants: The *Sinorhizobium* - *Medicago* model. *Nat Rev Microbiol* **5**: 619–633.

Joshi, P.K. and Rao, P.P. (2017) Global pulses scenario: status and outlook. *Ann N Y Acad Sci* **1392**: 6–17.

Joubès, J. and Chevalier, C. (2000) Endoreduplication in higher plants. *Plant Mol Biol* **43**: 735–745.

Kafri, R., Levy, M., and Pilpel, Y. (2006) The regulatory utilization of genetic redundancy through responsive backup circuits. *Proc Natl Acad Sci U S A* **103**: 11653–8.

Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., *et al.* (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res* **7**: 331–338.

Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., *et al.* (2002) Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res* **9**: 189–97.

Karmakar, K., Kundu, A., Rizvi, A.Z., Dubois, E., Severac, D., Czernic, P., *et al.* (2018) 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* MPMI-06-18-0174-R.

Karunakaran, R., Haag, A.F., East, A.K., Ramachandran, V.K., Prell, J., James, E.K., *et al.* (2010) BacA is essential for bacteroid development in nodules of Galegoid, but not Phaseoloid, legumes. *J Bacteriol* **192**: 2920–2928.

Karunakaran, R., Ramachandran, V.K., Seaman, J.C., East, A.K., Mouhsine, B., Mauchline, T.H., *et al.* (2009) Transcriptomic analysis of *Rhizobium leguminosarum* biovar *viciae* in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. *J Bacteriol* **191**: 4002–4014.

Kawaharada, Y., Kelly, S., Nielsen, M.W., Hjuler, C.T., Gysel, K., Muszyński, A., *et al.* (2015) Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature* **523**: 308–312.

Kazmierczak, T., Nagymihály, M., Lamouche, F., Barrière, 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.

Kiers, E.T., Rousseau, R.A., West, S.A., and Denison, R.F. (2003) Host sanctions and the legume–rhizobium mutualism. *Nature* **425**: 78–81.

Kim, M., Chen, Y., Xi, J., Waters, C., Chen, R., and Wang, D. (2015) An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proc Natl Acad Sci U S A* **112**: 15238–15243.

Kleemann, G., Alskog, G., Berry, A.M., and Huss-Danell, K. (1994) Lipid composition and nitrogenase activity of symbiotic *Frankia* (*Alnus incana*) in response to different oxygen concentrations. *Protoplasma* **183**: 107–115.

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.

Kondorosi, E., Mergaert, P., and Kereszt, A. (2013) A paradigm for endosymbiotic life: cell differentiation of rhizobium bacteria provoked by host plant factors. *Annu Rev Microbiol* **67**: 611–628.

Kumar, K., Mella-Herrera, R.A., and Golden, J.W. (2010) Cyanobacterial heterocysts. *Cold Spring Harb Perspect Biol* **2**: 1–19.

de la Fuente-Núñez, C., Mertens, J., Smit, J., and Hancock, R.E.W. (2012) The bacterial surface layer provides protection against antimicrobial peptides. *Appl Environ Microbiol* **78**: 5452–6.

Lancaster, K.M., Roemelt, M., Ettenhuber, P., Hu, Y., Ribbe, M.W., Neese, F., *et al.* (2011) X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. *Science* **334**: 974–977.

Lang, C. and Long, S.R. (2015) Transcriptomic analysis of *Sinorhizobium meliloti* and *Medicago truncatula* symbiosis using nitrogen fixation–deficient nodules. *Mol Plant-Microbe Interact* **28**: 856–868.

Langella, O., Hoogland, C., Joets, J., Zivy, M., Valot, B., Jacob, D., *et al.* (2013) Management and dissemination of MS proteomic data with PROTICdb: Example of a quantitative comparison between methods of protein extraction. *Proteomics* **13**: 1457–1466.

Lardi, M., Murset, V., Fischer, H.M., Mesa, S., Ahrens, C.H., Zamboni, N., and Pessi, G. (2016) Metabolomic profiling of *Bradyrhizobium diazoefficiens* -induced root nodules reveals both host plant-specific and developmental signatures. *Int J Mol Sci* **17**: 1–19.

Lavin, M., Pennington, R.T., Klitgaard, B.B., Sprent, J.I., de Lima, H.C., and Gasson, P.E. (2001) The dalbergioid legumes (Fabaceae): delimitation of a pantropical monophyletic clade. *Am J Bot* **88**: 503–533.

Lerouge, I. and Vanderleyden, J. (2001) O-antigen structural variation : mechanisms and possible roles in animal / plant-microbe interactions. *FEMS Microbiol Rev* **26**: 17–47.

Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J.C., and Dénarié, J. (1990) Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* **344**: 781–784.

LeVier, K., Phillips, R.W., Grippe, V.K., Roop, R.M., and Walker, G.C. (2000) Similar requirements of a plant symbiont and a mammalian pathogen for prolonged intracellular survival. *Science* **287**: 2492–3.

Liew, P.W.Y., Jong, B.C., and Najimudin, N. (2015) Hypothetical protein avin_16040 as the S-Layer protein of *Azotobacter vinelandii* and its involvement in plant root surface attachment. *Appl Environ Microbiol* **81**: 7484–7495.

Limpens, E., Mirabella, R., Fedorova, E., Franken, C., Franssen, H., Bisseling, T., and Geurts, R. (2005) Formation of organelle-like N2-fixing symbiosomes in legume root nodules is controlled by DMI2. *Proc Natl Acad Sci U S A* **102**: 10375–10380.

Ling, J., Wang, H., Wu, P., Li, T., Tang, Y., Naseer, N., *et al.* (2016) Plant nodulation inducers enhance horizontal gene transfer of *Azorhizobium caulinodans* symbiosis island. *Proc Natl Acad Sci U S A* **113**: 13875–13880.

Liu, C.-W. and Murray, J.D. (2016) The role of flavonoids in nodulation host-range specificity: an update. **5**: 33.

Loh, J. and Stacey, G. (2003) Nodulation gene regulation in Bradyrhizobium japonicum: a unique integration of global regulatory circuits. *Appl Environ Microbiol* **69**: 10 LP-17.

López-Torrejón, G., Jiménez-Vicente, E., Buesa, J.M., Hernandez, J.A., Verma, H.K., and Rubio, L.M. (2016) Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nat Commun* **7**: 1–6.

Maathuis, F.J. (2009) Physiological functions of mineral macronutrients. *Curr Opin Plant Biol* **12**: 250–258.

Manzoni, C., Kia, D.A., Vandrovcova, J., Hardy, J., Wood, N.W., Lewis, P.A., and Ferrari, R. (2018) Genome , transcriptome and proteome : the rise of omics data and their integration in biomedical sciences. **19**: 286–302.

Marchal, K. and Vanderleyden, J. (2000) The "oxygen paradox" of dinitrogen-fixing bacteria. *Biol Fertil Soils* **30**: 363–373.

Markmann, K. and Parniske, M. (2009) Evolution of root endosymbiosis with bacteria: how novel are nodules? *Trends Plant Sci* **14**: 77–86.

Martijn, J., Vosseberg, J., Guy, L., Offre, P., and Ettema, T.J.G. (2018) Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature* **557**: 101–105.

Maruya, J. and 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 Cell Physiol* **51**: 1443–1452.

Masson-Boivin, C., Giraud, E., Perret, X., and Batut, J. (2009) Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol* **17**: 458–466.

Masuda, T. and Goldsmith, P.D. (2009) World soybean production: Area harvested, yield, and long-term projections. *Int Food Agribus Manag Rev* **12**: 143–162.

McInerney, J.O., O'Connell, M.J., and Pisani, D. (2014) The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat Rev Microbiol* **12**: 449–455.

Meeks, J.C. and Elhai, J. (2002) Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiol Mol Biol Rev* **66**: 94–121.

Mergaert, P. (2018) Role of antimicrobial peptides in controlling symbiotic bacterial populations. *Nat Prod Rep* **35**: 336–356.

Mergaert, P., Kikuchi, Y., Shigenobu, S., and Nowack, E.C.M. (2017) Metabolic integration of bacterial endosymbionts through antimicrobial peptides. *Trends Microbiol* **25**: 703–712.

Mergaert, P., Van Montagu, M., and Holsters, M. (1997) Microreview: molecular mechanisms of Nod factor diversity. *Mol Microbiol* **25**: 811–817.

Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., and Kondorosi, E. (2003) A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiol* **132**: 161–73.

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.

Miller, A.J. and Cramer, M.D. (2004) Root nitrogen acquisition and assimilation. *Plant Soil* **274**: 1–36.

Minchin, F.R., James, E.K., and Becana, M. (2008) Oxygen diffusion, production of reactive oxygen and nitrogen species, and antioxidants in legume nodules. In, *Nitrogen-fixing leguminous symbioses*. Springer, pp. 321–362.

Molouba, F., Lorquin, J., Willems, A., Hoste, B., Giraud, E., Dreyfus, B., *et al.* (1999)

Photosynthetic bradyrhizobia from *Aeschynomene* spp. are specific to stem-nodulated species and form a separate 16S ribosomal DNA restriction fragment length polymorphism group. *Appl Environ Microbiol* **65**: 3084–3094.

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.

Mornico, D., Miché, L., Béna, G., Nouwen, N., Verméglio, A., Vallenet, D., *et al.* (2012) Comparative genomics of *Aeschynomene* symbionts: Insights into the ecological lifestyle of nod-independent photosynthetic bradyrhizobia. *Genes (Basel)* **3**: 35–61.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**: 621–8.

Moulin, L., Munive, A., Dreyfus, B., and Boivin-Masson, C. (2001) Nodulation of legumes by members of the beta-subclass of Proteobacteria. *Nature* **412**: 926-U17.

Mullineaux, C.W., Mariscal, V., Nenninger, A., Khanum, H., Herrero, A., Flores, E., and Adams, D.G. (2008) Mechanism of intercellular molecular exchange in heterocyst-forming cyanobacteria. *EMBO J* **27**: 1299–1308.

Muñoz-Amatriaín, M., Mirebrahim, H., Xu, P., Wanamaker, S.I., Luo, M.C., Alhakami, H., *et al.* (2017) Genome resources for climate-resilient cowpea, an essential crop for food security. *Plant J* **89**: 1042–1054.

Mus, F., Crook, M.B., Garcia, K., Garcia Costas, A., Geddes, B.A., Kouri, E.D., *et al.* (2016) Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. *Appl Environ Microbiol* **82**: 3698–3710.

Newton, W. (1998) Nitrogénases : fonction et évolution. In, *Bulletin Soc Fr de Microbiol*., pp. 238–241.

Nouwen, N., Arrighi, J.F., Cartieaux, F., Chaintreuil, C., Gully, D., Klopp, C., and Giraud, E. (2017) The role of rhizobial (NifV) and plant (FEN1) homocitrate synthases in *Aeschynomene*/photosynthetic *Bradyrhizobium* symbiosis. *Sci Rep* **7**: 1–10.

Okazaki, S., Kaneko, T., Sato, S., and Saeki, K. (2013) Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *Proc Natl Acad Sci U S A* **110**: 17131–17136.

Okazaki, S., Okabe, S., Higashi, M., Shimoda, Y., Sato, S., Tabata, S., *et al.* (2010) Identification and functional analysis of type III effector proteins in Mesorhizobium loti. *Mol Plant-Microbe Interact* **23**: 223–234.

Okazaki, S., Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., *et al.* (2015) Rhizobium–legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. *ISME J* **10**: 64–74.

Oke, V. and Long, S.R. (1999) Bacteroid formation in the Rhizobium–legume symbiosis. *Curr Opin Microbiol* **2**: 641–646.

Oldroyd, G.E.D. (2013) Speak, friend, and enter: Signalling systems that promote beneficial symbiotic associations in plants. *Nat Rev Microbiol* **11**: 252–263.

Oldroyd, G.E.D. and Dixon, R. (2014) Biotechnological solutions to the nitrogen problem. *Curr Opin Biotechnol* **26**: 19–24.

Oldroyd, G.E.D., Murray, J.D., Poole, P.S., and Downie, J.A. (2011) The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev Genet* **45**: 119–144.

Oono, R. and Denison, R.F. (2010) Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol* **154**: 1541–1548.

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**: 508–520.

Op den Camp, R.H.M., Polone, E., Fedorova, E., Roelofsen, W., Squartini, A., Op den Camp, H.J.M., *et al.* (2012) Nonlegume *Parasponia andersonii* deploys a broad rhizobium host range strategy resulting in largely variable symbiotic effectiveness. *Mol Plant-Microbe Interact* **25**: 954–963.

Oshlack, A. and Wakefield, M.J. (2009) Transcript length bias in RNA-seq data confounds systems biology. *Biol Direct* **4**: 14.

Penterman, J., Abo, R.P., De Nisco, N.J., Arnold, M.F.F., Longhi, R., Zanda, M., and Walker, G.C. (2014) Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. *Proc Natl Acad Sci U S A* **111**: 3561–3566.

Pérez-Montaño, F., del Cerro, P., Jiménez-Guerrero, I., López-Baena, F.J., Cubo, M.T., Hungria, M., *et al.* (2016) RNA-seq analysis of the *Rhizobium tropici* CIAT 899 transcriptome shows similarities in the activation patterns of symbiotic genes in the presence of apigenin and salt. *BMC Genomics* **17**: 1–11.

Pérez-Montaño, F., Jiménez-Guerrero, I., Acosta-Jurado, S., Navarro-Gómez, P., Ollero, F.J., Ruiz-Sainz, J.E., *et al.* (2016) A transcriptomic analysis of the effect of genistein on *Sinorhizobium fredii* HH103 reveals novel rhizobial genes putatively involved in symbiosis. *Sci Rep* **6**: 1–12.

Perret, X., Staehelin, C., and Broughton, W.J. (2000) Molecular basis of symbiotic promiscuity. *Microbiol Mol Biol Rev* **64**: 180–201.

Pessi, G., Ahrens, C.H., Rehrauer, H., Lindemann, A., Hauser, F., Fischer, H.-M., and Hennecke, H. (2007) Genome-wide transcript analysis of *Bradyrhizobium japonicum* bacteroids in soybean root nodules. *Mol Plant-Microbe Interact* **20**: 1353–1363.

Peters, N.K., Frost, J.W., and Long, S.R. (1986) A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* **233**: 977–980.

Petersen, J.M., Kemper, A., Gruber-Vodicka, H., Cardini, U., Van Der Geest, M., Kleiner, M., *et al.* (2016) Chemosynthetic symbionts of marine invertebrate animals are capable of nitrogen fixation. *Nat Microbiol* **2**: 1–11.

Pini, F., Frage, B., Ferri, L., De Nisco, N.J., Mohapatra, S.S., Taddei, L., *et al.* (2013) The DivJ, CbrA and PleC system controls DivK phosphorylation and symbiosis in *Sinorhizobium meliloti*. *Mol Microbiol* **90**: 54–71.

Pini, F., De Nisco, N.J., Ferri, L., Penterman, J., Fioravanti, A., Brilli, M., *et al.* (2015) Cell cycle control by the master regulator CtrA in *Sinorhizobium meliloti*. *PLoS Genet* **11**: 1–24.

Poole, P., Ramachandran, V., and Terpolilli, J. (2018) Rhizobia: from saprophytes to endosymbionts. *Nat Rev Microbiol* **16**: 291–303.

Prell, J., White, J.P., Bourdes, A., Bunnewell, S., Bongaerts, R.J., and Poole, P.S. (2009) Legumes regulate Rhizobium bacteroid development and persistence by the supply of branched-chain amino acids. *Proc Natl Acad Sci U S A* **106**: 12477–82.

Price, M.N., Deutschbauer, A.M., Skerker, J.M., Wetmore, K.M., Ruths, T., Mar, J.S., *et al.* (2013) Indirect and suboptimal control of gene expression is widespread in bacteria. *Mol*

Syst Biol **9**: 660.

Price, P.A., Tanner, H.R., Dillon, B.A., Shabab, M., Walker, G.C., and Griffitts, 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.

Pueppke, S.G. and Broughton, W.J. (1999) *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. *Mol Plant-Microbe Interact* **12**: 293–318.

Rai, A.N., Soderback, E., and Bergman, B. (2000) Tansley Review No. 116. Cyanobacterium-Plant Symbioses. *New Phytol* **147**: 449–481.

Ramaswamy, P. and Bal, A.K. (1986) Media-induced changes in the asymbiotic nitrogen-fixing bacteroids of *Bradyrhizobium* sp. 32H1. *Curr Microbiol* **14**: 181–185.

Rastogi, R.P., Madamwar, D., and Incharoensakdi, A. (2015) Bloom dynamics of cyanobacteria and their toxins: environmental health impacts and mitigation strategies. *Front Microbiol* **6**: 1–22.

Rawat, N., Kumari, K., Singh, F., and Gilhare, V.R. (2015) Effect of *Azolla* -supplemented feeding on milk production of cattle and production performance of broilers. *Appl Biol Res* **17**: 214.

Raymond, J., Siefert, J.L., Staples, C.R., and Blankenship, R.E. (2004) The Natural History of Nitrogen Fixation. *Mol Biol Evol* **21**: 541–554.

Remigi, P., Capela, D., Clerissi, C., Tasse, L., Torchet, R., Bouchez, O., *et al.* (2014) Transient hypermutagenesis accelerates the evolution of legume endosymbionts following horizontal gene transfer. *PLoS Biol* **12**: e1001942.

Remigi, P., Zhu, J., Young, J.P.W., and Masson-Boivin, C. (2016) Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. *Trends Microbiol* **24**: 63–75.

Renier, A., Maillet, F., Fardoux, J., Poinsot, 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.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**: e47–e47.

Robledo, M., Jimenez-Zurdo, J.I., Velazquez, E., Trujillo, M.E., Zurdo-Pineiro, J.L., Ramirez-Bahena, M.H., *et al.* (2008) Rhizobium cellulase CelC2 is essential for primary symbiotic infection of legume host roots. *Proc Natl Acad Sci U S A* **105**: 7064–7069.

Rogel, M.A., Ormeño-Orrillo, E., and Martinez Romero, E. (2011) Symbiovars in rhizobia reflect bacterial adaptation to legumes. *Syst Appl Microbiol* **34**: 96–104.

Rogers, C. and Oldroyd, G.E.D. (2014) Synthetic biology approaches to engineering the nitrogen symbiosis in cereals. *J Exp Bot* **65**: 1939–46.

Rohart, F., Gautier, B., Singh, A., and Lê Cao, K.-A. (2017) mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol* **13**: e1005752.

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**: 817–837.

Rubio, L.M. and Ludden, P.W. (2008) Biosynthesis of the iron-molybdenum cofactor of

nitrogenase. *Annu Rev Microbiol* **62**: 93–111.

Sagan, L. (1967) On the origin of mitosing cells. *J Theor Biol* **14**: 225-274, IN1-IN6.

Sallet, E., Gouzy, J., and Schiex, T. (2014) EuGene-PP: a next-generation automated annotation pipeline for prokaryotic genomes. **30**: 2659–2661.

Santi, C., Bogusz, D., and Franche, C. (2013) Biological nitrogen fixation in non-legume plants. *Ann Bot* **111**: 743–767.

Schmutz, J., McClean, P.E., Mamidi, S., Wu, G.A., Cannon, S.B., Grimwood, J., *et al.* (2014) A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet* **46**: 707–713.

Seefeldt, L.C., Hoffman, B.M., and Dean, D.R. (2009) Mechanism of Mo-dependent nitrogenase. *Annu Rev Biochem* **78**: 701–722.

Sen, D. and 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 Sci Lett* **34**: 239–246.

Setten, L., Soto, G., Mozzicafreddo, M., Fox, A.R., Lisi, C., Cuccioloni, M., *et al.* (2013) Engineering *Pseudomonas protegens* Pf-5 for nitrogen fixation and its application to improve plant growth under nitrogen-deficient conditions. *PLoS One* **8**: e63666.

Silipo, A., Leone, M.R., Erbs, G., Lanzetta, R., Parrilli, M., Chang, W.S., *et al.* (2011) A unique bicyclic monosaccharide from the *Bradyrhizobium* lipopolysaccharide and its role in the molecular interaction with plants. *Angew Chemie - Int Ed* **50**: 12610–12612.

Silipo, A., Vitiello, G., Gully, D., Sturiale, L., Chaintreuil, C., Fardoux, J., *et al.* (2014) Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. *Nat Commun* **5**:.

Simonsen, A.K., Dinnage, R., Barrett, L.G., Prober, S.M., and Thrall, P.H. (2017) Symbiosis limits establishment of legumes outside their native range at a global scale. *Nat Commun* **8**: 1–9.

Sipe, A.R., Wilbur, A.E., and Cary, S.C. (2000) Bacterial symbiont transmission in the wood-boring shipworm *Bankia setacea* (Bivalvia: Teredinidae). *Appl Environ Microbiol* **66**: 1685–1691.

Smil, V. (2001) Enriching the Earth. Fritz Haber, Carl Bosch, and the transfromation of world food production. *MIT Press Cambridge, Massachusetts* 358 p.

Sprent, J.I., Ardley, J., and James, E.K. (2017) Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytol* **215**: 40–56.

Sprent, J.I., Ardley, J.K., and James, E.K. (2013) From North to South: A latitudinal look at legume nodulation processes. *South African J Bot* **89**: 31–41.

Sprent, J.I. and Gehlot, H.S. (2010) Nodulated legumes in arid and semi-arid environments: Are they important? *Plant Ecol Divers* **3**: 211–219.

Starker, C.G., Parra-Colmenares, A.L., Smith, L., Mitra, R.M., and Long, S.R. (2006) Nitrogen fixation mutants of *Medicago truncatula* fail to support plant and bacterial symbiotic gene expression. *Plant Physiol* **140**: 671–80.

Steffen, W., Richardson, K., Rockstrom, J., Cornell, S.E., Fetzer, I., Bennett, E.M., *et al.* (2015) Planetary boundaries: Guiding human development on a changing planet. *Science* **347**: 1259855–1259855.

Strous, M., Fuerst, J.A., Kramer, E.H.M., Logemann, G.M., van de Pas-Schoonen, K.T., Webb, R., *et al.* (1999) Missing lithotroph identifed as new planctomycete. *Nature* **400**: 446–449.

Strullu-Derrien, C., Selosse, M.-A., Kenrick, P., and Martin, F.M. (2018) The origin and evolution of mycorrhizal symbioses: From palaeomycology to phylogenomics. *New Phytol* **220**: 1012–1030.

Svistoonoff, S., Hocher, V., and Gherbi, H. (2014) Actinorhizal root nodule symbioses: What is signalling telling on the origins of nodulation? *Curr Opin Plant Biol* **20**: 11–18.

Taulé, C., Mareque, C., Barlocco, C., Hackembruch, F., Reis, V.M., Sicardi, M., and Battistoni, F. (2012) The contribution of nitrogen fixation to sugarcane (*Saccharum officinarum* L.), and the identification and characterization of part of the associated diazotrophic bacterial community. *Plant Soil* **356**: 35–49.

Terpolilli, J.J., Hood, G.A., and Poole, P.S. (2012) What determines the efficiency of N2-fixing rhizobium-legume symbioses? In, *Advances in Microbial Physiology*. Elsevier Ltd., pp. 325–389.

Thilakarathna, M.S. and Raizada, M.N. (2017) A meta-analysis of the effectiveness of diverse rhizobia inoculants on soybean traits under field conditions. *Soil Biol Biochem* **105**: 177–196.

Thomas, J. (1970) Absence of the pigments of photosystem II of photosynthesis in heterocysts of a blue-green alga. *Nature* **228**: 181–183.

Timmers, A.C.J., Soupène, E., Auriac, M.-C., de Billy, F., Vasse, J., Boistard, P., and Truchet, G. (2000) Saprophytic intracellular rhizobia in alfalfa nodules. *Mol Plant-Microbe Interact* **13**: 1204–1213.

Tiricz, H., Szücs, 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.

Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2011) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol* **10**: 123–36.

Udvardi, M. and Poole, P.S. (2013) Transport and Metabolism in Legume-Rhizobia Symbioses. *Annu Rev Plant Biol* **64**: 781–805.

Valladares, A., Herrero, A., Pils, D., Schmetterer, G., and Flores, E. (2003) Cytochrome c oxidase genes required for nitrogenase activity and diazotrophic growth in *Anabaena* sp . PCC 7120. **47**: 1239–1249.

Vallenet, D., Calteau, A., Cruveiller, S., Gachet, M., Lajus, A., Josso, A., *et al.* (2017) MicroScope in 2017: An expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res* **45**: D517–D528.

Valot, B., Langella, O., Nano, E., and Zivy, M. (2011) MassChroQ: A versatile tool for mass spectrometry quantification. *Proteomics* **11**: 3572–3577.

Vasse, J., De Billy, F., Camut, S., and Truchet, G. (1990) Correlation between ultrastructural differentiation of bacteriods and nitrogen fixation in alfalfa nodules. *J Bacteriol* **172**: 4295–4306.

Vauclare, P., Bligny, R., Gout, E., and Widmer, F. (2013) An overview of the metabolic differences between *Bradyrhizobium japonicum* 110 bacteria and differentiated bacteroids from soybean (*Glycine max*) root nodules: An in vitro 13C- and 31P-nuclear magnetic resonance spectroscopy study. *FEMS Microbiol Lett* **343**: 49–56.

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.

van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., *et al.* (2018) Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. *Proc Natl Acad Sci U S A* **115**: E4700–E4709.

Venkateshwaran, M., Jayaraman, D., Chabaud, M., Genre, A., Balloon, A.J., Maeda, J., *et al.* (2015) A role for the mevalonate pathway in early plant symbiotic signaling. *Proc Natl Acad Sci U S A* **112**: 9781–9786.

Villareal, T.A. (1992) Marine nitrogen-fixing diatom-cyanobacteria symbioses. In, *Marine Pelagic Cyanobacteria: Trichodesmium and other Diazotrophs*. Springer Netherlands, Dordrecht, pp. 163–175.

Vinardell, J.M., Fedorova, E., Cebolla, A., Kevei, Z., Horvath, G., Kelemen, Z., *et al.* (2003) Endoreduplication mediated by the anaphase-promoting complex activator CCS52A is required for symbiotic cell differentiation in *Medicago truncatula* nodules. *Plant Cell* **15**: 2093–105.

Voisin, A.-S., Guéguen, J., Huyghe, C., Jeuffroy, M.-H., Magrini, M.-B., Meynard, J.M., *et al.* (2013) Les légumineuses dans l'Europe du XXIè siècle : Quelle place dans les systèmes agricoles et alimentaires actuels et futurs ? Quels nouveaux défis pour la recherche ? *Innov Agron* **30**: 283–312.

Wang, D., Griffitts, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., *et al.* (2010) A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* **327**: 1126–1129.

Wang, L., Zhang, L., Liu, Z., Liu, Z., Zhao, D., Liu, X., *et al.* (2013) A minimal nitrogen fixation gene cluster from *Paenibacillus* sp. WLY78 enables expression of active nitrogenase in *Escherichia coli*. *PLoS Genet* **9**: e1003865.

Wang, Q., Liu, J., Li, H., Yang, S., Körmöczi, P., Kereszt, A., and Zhu, H. (2017) Nodule-specific cysteine-rich peptides negatively regulate nitrogen-fixing symbiosis in a strain-specific manner in Medicago truncatula. *Mol Plant-Microbe Interact* **31**: 240–248.

Wang, Q., Yang, S., Liu, J., Terecskei, K., Ábrahám, E., Gombár, A., *et al.* (2017) Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. *Proc Natl Acad Sci U S A* **114**: 201700715.

Ward, B. (2012) The Global Nitrogen Cycle. In, *Fundamentals of Geobiology*. John Wiley & Sons, Ltd, Chichester, UK, pp. 36–48.

Wehmeier, S., Arnold, M.F.F., Marlow, V.L., Aouida, M., Myka, K.K., Fletcher, V., *et al.* (2010) Internalization of a thiazole-modified peptide in *Sinorhizobium meliloti* occurs by BacA-dependent and -independent mechanisms. **156**: 2702–2713.

Williams, T.A., Foster, P.G., Cox, C.J., and Embley, T.M. (2013) An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* **504**: 231–236.

Wojciechowski, M.F., Sanderson, M.J., Steele, K.P., and Liston, a. (2000) Molecular phylogeny of the "Temperate Herbaceous Tribes" of papilionoid legumes: A supertree approach. *Adv Legum Syst* **9**: 277–298.

Wong, P.P. and Burris, R.H. (1972) Nature of oxygen inhibition of nitrogenase from Azotobacter vinelandii. *Proc Natl Acad Sci U S A* **69**: 672–675.

Xie, F., Murray, J.D., Kim, J., Heckmann, A.B., Edwards, A., Oldroyd, G.E.D., and Downie, J.A. (2012) Legume pectate lyase required for root infection by rhizobia. **109**: 633–638.

Yang, S., Wang, Q., Fedorova, E., Liu, J., Qin, Q., Zheng, Q., *et al.* (2017) Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proc Natl Acad Sci U S A* **114**: 201700460.

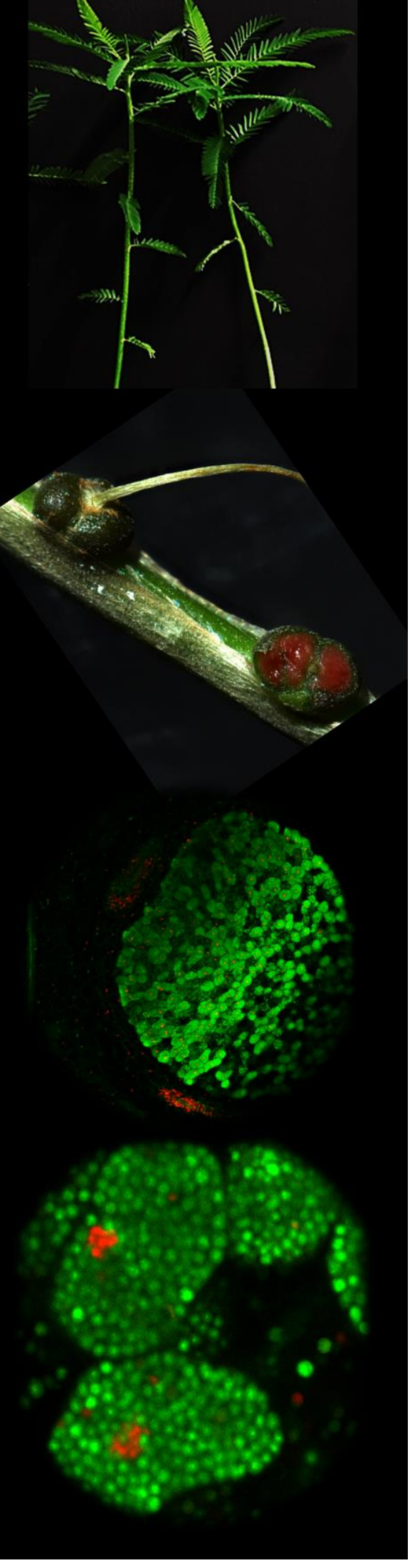
Young, N.D., Debellé, F., Oldroyd, G.E.D., Geurts, R., Cannon, S.B., Udvardi, M.K., *et al.* (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature* **480**: 520–524.

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.

Zhao, L., Deng, Z., Yang, W., Cao, Y., Wang, E., and Wei, G. (2010) Diverse rhizobia associated with *Sophora alopecuroides* grown in different regions of Loess Plateau in China. *Syst Appl Microbiol* **33**: 468–477.

Zimorski, V., Ku, C., Martin, W.F., and Gould, S.B. (2014) Endosymbiotic theory for organelle origins. *Curr Opin Microbiol* **22**: 38–48.

Zipfel, C. and Oldroyd, G.E.D. (2017) Plant signalling in symbiosis and immunity. *Nature* **543**: 328–336.



Annexes

Annexes

1. Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium* strains

Théophile Kazmierczak, Marianna Nagymihály, Florian Lamouche, Quentin Barrière, Ibtissem Guefrachi, Benoit Alunni, Mouna Ouadghiri, Jamal Ibijbijen, Éva Kondorosi, Peter Mergaert et Véronique Gruber

Publié dans la revue *Molecular Plant Microbe Interactions* (2017).

DOI: https://doi.org/10.1094/MPMI-01-17-0009-R

MPMI Vol. 30, No. 5, 2017, pp. 399–409. http://dx.doi.org/10.1094/MPMI-01-17-0009-R

e-Xtra*

Specific Host-Responsive Associations Between *Medicago truncatula* Accessions and *Sinorhizobium* Strains

Théophile Kazmierczak,[1] Marianna Nagymihály,[2,3] Florian Lamouche,[2] Quentin Barrière,[2] Ibtissem Guefrachi,[2] Benoit Alunni,[2] Mouna Ouadghiri,[4] Jamal Ibijbijen,[5] Éva Kondorosi,[2,3] Peter Mergaert,[2] and Véronique Gruber[1]

[1]Institute of Plant Sciences Paris Saclay IPS2, CNRS, INRA, Université Paris-Sud, Université Evry, Université Paris-Saclay, Bâtiment 630, 91405 Orsay, France. Institute of Plant Sciences Paris-Saclay IPS2, Université Paris Diderot, Sorbonne Paris-Cité, Bâtiment 630, 91405, Orsay, France; [2]Institute for Integrative Biology of the Cell, UMR 9198, CNRS/Université Paris-Sud/CEA, 91198, Gif-sur-Yvette, France; [3]Institute of Biochemistry, Biological Research Centre, Hungarian Academy of Sciences, 6726 Szeged, Hungary; [4]Collections Coordonnées Marocaines de Microorganismes et Laboratoire de Microbiologie et Biologie Moléculaire, Centre National pour la Recherche Scientifique et Technique, Rabat, Maroc; and [5]Faculté des Sciences, Université Moulay Ismail, Meknès, Maroc, Faculté des Sciences, Université Moulay Ismail, BP 11201 Zitoune, Meknès, Maroc

Accepted 7 March 2017.

Legume plants interact with rhizobia to form nitrogen-fixing root nodules. Legume-rhizobium interactions are specific and only compatible rhizobia and plant species will lead to nodule formation. Even within compatible interactions, the genotype of both the plant and the bacterial symbiont will impact on the efficiency of nodule functioning and nitrogen-fixation activity. The model legume *Medicago truncatula* forms nodules with several species of the *Sinorhizobium* genus. However, the efficiency of these bacterial strains is highly variable. In this study, we compared the symbiotic efficiency of *Sinorhizobium meliloti* strains Sm1021, 102F34, and FSM-MA, and *Sinorhizobium medicae* strain WSM419 on the two widely used *M. truncatula* accessions A17 and R108. The efficiency of the interactions was determined by multiple parameters. We found a high effectiveness of the FSM-MA strain with both *M. truncatula* accessions. In contrast, specific highly efficient interactions were obtained for the A17-WSM419 and R108-102F34 combinations. Remarkably, the widely used Sm1021 strain performed weakly on both hosts. We showed that Sm1021 efficiently induced nodule organogenesis but cannot fully activate the differentiation of the symbiotic nodule cells, explaining its weaker performance. These results will be informative for the selection of appropriate rhizobium strains in functional studies on symbiosis using these *M. truncatula* accessions, particularly for research focusing on late stages of the nodulation process.

Soil bacteria called rhizobia interact with the root system of legume plants in a coordinated developmental program to form symbiotic root nodules. Within the cells of these nodules, nitrogen-fixing rhizobia provide the host plant with fixed nitrogen. The rhizobium-legume symbiosis attracts considerable interest because of its impact on the global nitrogen cycle and its potential biotechnological application for nitrogen nutrition in agriculture (Oldroyd and Dixon 2014). *Medicago truncatula* is a legume species that has been intensively studied because of its attractive features as a laboratory model system (Cheng et al. 2014; Rose 2008).

M. truncatula forms indeterminate nodules characterized by distinct developmental zones, organized from the youngest part at the nodule tip to the oldest root proximal region (Ferguson et al. 2010; Kondorosi et al. 2013; Lotocka et al. 2012; Oldroyd and Downie 2008). A persistent apical meristem (zone I) is composed of dividing cells and is responsible for the indeterminate growth of the nodule. In the infection/differentiation zone (zone II) postmeristematic cells are invaded by rhizobia and the infected nodule cells differentiate along with their intracellular rhizobia. In the nitrogen-fixation zone (zone III), the differentiated bacteria, bacteroids, and their host symbiotic nodule cells are mature and fix atmospheric nitrogen. Aging nodule cells constitute the senescence zone (zone IV) in which first the bacteroids and, then, the host cells degenerate.

The differentiation of the nodule cells in zone II is marked by a dramatic increase in their cell size, which is driven by several endoreduplication (ER) cycles leading to mature symbiotic nodule cells with a nuclear DNA content up to 64C (C is the haploid DNA content) and a size, about 50 µm in diameter, that is up to 80-fold larger than their progenitor cells in the meristem (Vinardell et al. 2003). The polyploidy of the symbiotic nodule cells is essential to their function, since knockdown of *CCS52A*, a regulator of the ER step, leads to nonfunctional nodules (Vinardell et al. 2003). Moreover, the formation of polyploid nodule cells is correlated with the activation of expression of hundreds of nodule-specific genes (Maunoury et al. 2010; Roux et al. 2014).

The rhizobia that are released into the cells of zone II also differentiate concomitantly with their host cells (Mergaert et al. 2006; Vasse et al. 1990). The released rhizobia first proliferate, partially filling the infected cell. However, their division is rapidly inhibited. The bacteria, nevertheless, continue to grow by cell enlargement driven by polyploidization, not unlike the

T. Kazmierczak and M. Nagymihály contributed equally to this work.

Current address for I. Guefrachi: Université de Pau et des Pays de l'Adour, Pau, France.

Corresponding authors: P. Mergaert; E-mail: peter.mergaert@i2bc.paris-saclay.fr and V. Gruber; E-mail: veronique.gruber@ips2.universite-paris-saclay.fr

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

© 2017 The American Phytopathological Society

growth of their host cells. At the end of the process, the bacteria entirely fill the host cell and are terminally differentiated into large polyploid nitrogen-fixing bacteroids. The process of terminal differentiation of bacteroids is mediated by host cell-produced peptides known as nodule-specific cysteine-rich (NCR) peptides (Alunni and Gourion 2016; Penterman et al. 2014; Van de Velde et al. 2010).

The *M. truncatula* subsp. *tricycla* R108 (Hoffmann et al. 1997; Yoder et al. 2013) and *M. truncatula* Jemalong A17 (Young et al. 2011) accessions are particularly widely used for research on the legume-rhizobium symbiosis because of their superior transformation capacity and the availability of extensive transcriptome data, genome sequence, and large-scale mutant collections (Cheng et al. 2014; Young et al. 2011; the *Medicago truncatula* Hapmap Project database). These plants establish symbiosis with several species of the *Sinorhizobium* genus, including *S. meliloti* and *S. medicae* strains, although the symbiotic efficiency (i.e., the plant benefit derived from the symbiosis) of these strains is variable (Laguerre et al. 2012; Larrainzar et al. 2014; Liu et al. 2014). *S. meliloti* is among the most characterized and first sequenced rhizobia (Galibert et al. 2001) and, therefore, the reference strain Sm1021, its sister strain Sm2011, strain Rm41, and their derived mutants are largely used in the *Medicago* research community. However, these *S. meliloti* strains, which were originally isolated from *Medicago sativa* nodules, are poor symbiotic partners for Jemalong A17 (Terpolilli et al. 2008). These strains efficiently induce nodules but these nodules are not or are only weakly functional and display low nitrogen-fixation activity. Although this is not necessarily a shortcoming in studies that focus on the very early stages of the interaction, their use is problematic in research on the late stages of symbiosis. Therefore, the strain *S. medicae* WSM419, which was isolated from *Medicago murex* and seemed to perform better on *M. truncatula* A17, has been proposed as an alternative to Sm1021 (Laguerre et al. 2012; Larrainzar et al. 2014; Liu et al. 2014; Reeve et al. 2010; Terpolilli et al. 2008). On the other hand, no comparative studies of the symbiotic performance of *Sinorhizobium* strains on *M. truncatula* subsp. *tricycla* R108 are available. Therefore, there is a need to identify appropriate bacterial strains interacting efficiently with the two *M. truncatula* accessions Jemalong A17 and R108. In this study, we compared the symbiotic efficiency of four different strains belonging to the two closely related species *S. meliloti* and *S. medicae* on the two *M. truncatula* accessions, using a combination of physiological, histological, and molecular analyses.

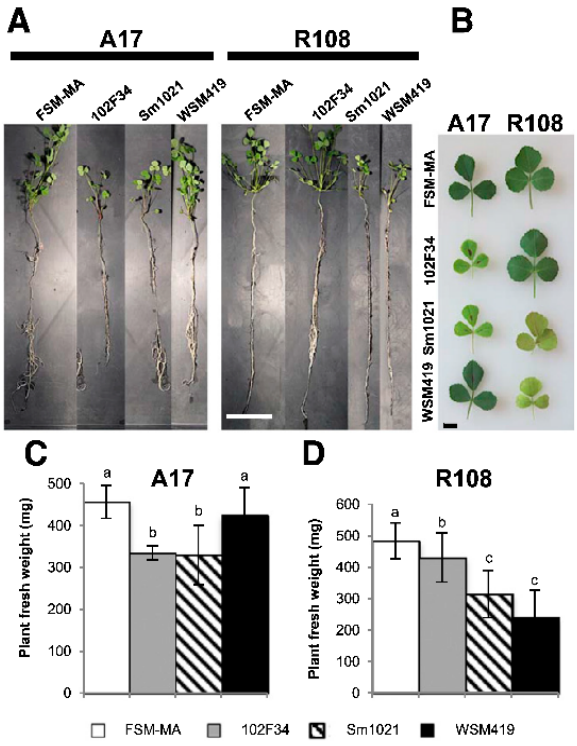


Fig. 1. Plant growth of *Medicago truncatula* A17 and R108 inoculated with different *Sinorhizobium* strains. **A,** Phenotypes of *M. truncatula* A17 and R108 at 5 weeks postinoculation (wpi) with bacterial strains FSM-MA, 102F34, Sm1021, and WSM419. All the plants were photographed at the same scale. The scale bar represents 10 cm. **B,** Leaf phenotypes of *M. truncatula* A17 and R108 at 5 wpi with the different bacterial strains FSM-MA, 102F34, Sm1021, and WSM419. First foliar rank leaves were pictured. Scale bars represent 0.5 cm. **C,** Plant biomass production (fresh weight) of *Medicago truncatula* A17 at 5 wpi with bacterial strains FSM-MA, 102F34, Sm1021, and WSM419. Values represent the mean ± standard error (n = 10). Letters indicate statistical differences between groups (P < 0.05, Kruskal-Wallis test). **D,** Plant biomass production (fresh weight) of *M. truncatula* R108 at 5 wpi with the bacterial strains FSM-MA, 102F34, Sm1021, and WSM419. Values represent the mean ± standard error (n = 10). Letters indicate statistical differences between groups (P < 0.05, Kruskal-Wallis test).

RESULTS

Plant growth phenotypes and nitrogenase activities reflect specific interactions between *M. truncatula* accessions and *Sinorhizobium* strains.

As a first approach to investigate the effectiveness of the four *Sinorhizobium* strains, *S. meliloti* FSM-MA, *S. meliloti* 102F34, *S. meliloti* Sm1021, and *S. medicae* WSM419, plant morphological responses to these strains were assessed in the two *M. truncatula* accessions Jemalong A17 and R108. The macroscopic phenotypes of five-week-old plants differed between the two *Medicago* accessions in interaction with the four bacterial strains (Fig. 1A; Supplementary Fig. 1). A17 plants had a higher biomass when inoculated with strains FSM-MA and WSM419 than with the other strains (Fig. 1A and C), whereas R108 plants had a higher biomass when inoculated with FSM-MA and 102F34 than with the other strains (Fig. 1A and D).

In addition, the efficiency of the different plant-rhizobium partnerships could be visualized by the degree of leaf chlorosis, which reflects the level of nitrogen starvation. In the A17 accession, plants inoculated with 102F34 or Sm1021 showed pale green, yellowing leaves in contrast to those infected with FSM-MA or WSM419, which had dark green leaves (Fig. 1B). In the R108 accession, leaf yellowing was observed for plants inoculated with Sm1021 and WSM419, whereas the leaves of plants inoculated with strains FSM-MA and 102F34 were dark green (Fig. 1B).

Next, we monitored nitrogen-fixation activity using the acetylene reduction assay (ARA) at 5 weeks postinoculation (wpi) on *M. truncatula* Jemalong A17 or R108 plants infected with each of the four rhizobial strains. For A17, nitrogenase activity was significantly greater for plants inoculated with FSM-MA and WSM419 than with 102F34 and Sm1021 (Fig. 2A). For R108, the highest nitrogenase activities were observed in plants inoculated with FSM-MA and 102F34 (Fig. 2A). Similar patterns were obtained for total nitrogenase activity per plant and the specific nitrogenase activity per nodule mass, suggesting variations between the strains and hosts in the functioning of the nodules (Fig. 2B).

Together, these results highlight a different ability of the tested *Sinorhizobium* strains to support plant development and, inversely, the two host accessions respond differently to the panel of tested strains. To extend our observations on *M. truncatula*, we

also analyzed the performance of these four *Sinorhizobium* strains on three different *Medicago sativa* cultivars and found that the FSM-MA, Sm1021, and 102F34 strains were very efficient on the three cultivars but the WSM419 strain was a weak symbiont for one of the cultivars (Supplementary Fig. 2; data not shown). Thus, taken together, the FSM-MA strain seems to have, among the four tested strains, the broadest host range for optimal symbiotic performance on *Medicago* species.

The symbiotic efficiency of the interactions between *M. truncatula* accessions and *Sinorhizobium* strains are correlated with nodule developmental parameters.

To explore the differences in the efficiency of symbiotic couples, we analyzed, in detail, the phenotypes of their nodules. Overall, the four strains formed equal numbers of nodules on the two plant accessions, except for strain 102F34 on *M. truncatula* A17 inducing almost twice as many nodules as the three other strains (Supplementary Fig. 3). In addition, small variations were observed in the total nodule biomass. A higher nodule biomass was observed in A17 plants inoculated with FSM-MA and WSM419 and, in R108, plants inoculated with FSM-MA and 102F34. Together, the quantification of nodule numbers and mass suggests that the differences between the strains are not explained by differences in the induction of nodules but are, rather, related to differences in nodule development.

An easy-to-score marker of nodule functioning is their color. Indeed, well-functioning nodules are pink because of the high production of leghemoglobin protein. This oxygen scavenger prevents inactivation of the oxygen-sensitive nitrogenase by maintaining a low free-oxygen concentration in the cytoplasm of the infected cells. At 3 wpi, the nodules were pink on A17 roots infected with the FSM-MA and WSM419 strains and on R108 roots inoculated with the FSM-MA and 102F34 strains, in agreement with the effective nitrogen-fixation activity in these nodules. On the contrary, the nodules of the other host plant–*Sinorhizobium* strain combinations were only faintly pink, correlating with the low nitrogenase activity in these associations.

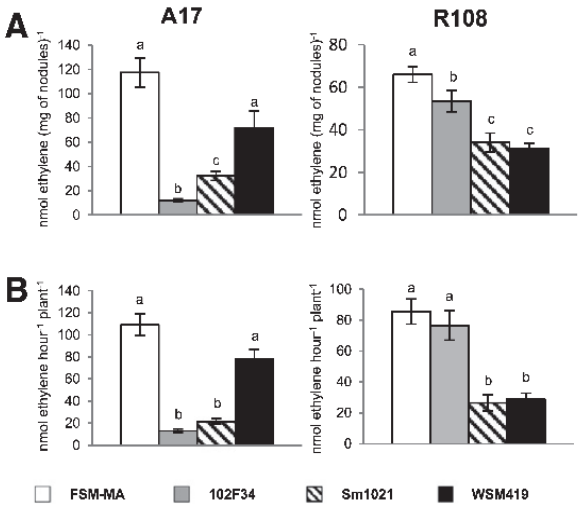


Fig. 2. Nitrogen-fixation efficiency of nodules of *Medicago truncatula* A17 and R108 inoculated with different *Sinorhizobium* strains. **A,** Specific nitrogenase activity (nmol ethylene per hour per milligram of dry nodules) of *M. truncatula* A17 and R108 at 5 weeks postinoculation (wpi) with bacterial strains FSM-MA, 102F34, Sm1021, and WSM419. **B,** Total nitrogenase activity (nmol ethylene per hour per plant) of *M. truncatula* A17 and R108 at 5 wpi with bacterial strains FSM-MA, 102F34, Sm1021, and WSM419. Values represent the mean ± standard error (n = 10). Letters indicate statistical differences between groups ($P < 0.05$, Kruskal-Wallis test).

To test whether the defect in symbiotic nitrogen fixation was related to nitrogenase expression, we examined the promoter activity of the nitrogenase reductase *nifH* gene, using a *nifH::GUS* (β-glucuronidase) reporter fusion. The *nifH::GUS* expression was induced under symbiotic conditions in all the tested *M. truncatula–Sinorhizobium* combinations (Supplementary Fig. 4). Although chromogenic staining for GUS activity is not quantitative, this observation demonstrates that, in all combinations, the nitrogenase gene expression is properly activated.

To investigate the histological organization of the nodules in the eight symbiotic couples, nodule sections were stained with the Live/Dead BacLight probe (Fig. 3), which is a mixture of two fluorescent DNA intercalating agents, SYTO9 and propidium iodide (PI). With this staining procedure, living rhizobia are stained with SYTO9 and are green fluorescent whereas dying cells are stained in red with PI. Also the nuclei of the host cells stain with PI in nodule sections (Haag et al. 2011). Zone III of A17 nodules infected with the FSM-MA and WSM419 strains and those of R108 induced by the FSM-MA and 102F34 strains were nearly entirely filled with nodule cells occupied by bacteroids, indicating the invasion of the whole root nodule with viable bacteroids (Fig. 3). In contrast, only a few cell layers full of viable bacteroids were detected in A17 nodules infected with the strains 102F34 and Sm1021 or in those of R108 infected with Sm1021 and WSM419 (Fig. 3). Thus the degree of the occupation of the nodules with viable bacteroid-containing symbiotic nodule cells correlates with the activity of the nodules.

To further investigate the symbiotic nodule cell formation, we measured the ploidy levels of nuclei from nodules harvested from A17 and R108 plants infected with the four studied rhizobia strains, using flow cytometry (Fig. 4). *M. truncatula* nodules contain nuclei with ploidy levels from 2C to 64C

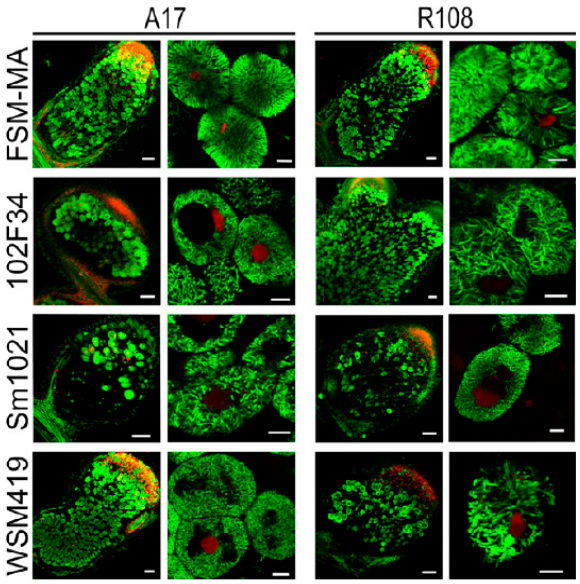


Fig. 3. Histological organization of nodules from *Medicago truncatula* A17 and R108 inoculated with different *Sinorhizobium* strains. Nodule sections were stained with the Live/Dead BacLight probe marking viable bacteroids with green fluorescence from SYTO9 and dead bacteria or plant nuclei in red fluorescence by propidium iodide. Representative confocal microscopy pictures of whole nodules at 5 weeks postinoculation or of infected cells in zone III of nodules isolated from A17 and R108 plants inoculated with strains FSM-MA, 102F34, Sm1021, and WSM419. The scale bars represent 100 µm (whole nodules) or 10 µm (infected cells).

(Maunoury et al. 2010; Vinardell et al. 2003). The highly endoreduplicated (polyploid) nuclei with 16C, 32C, and 64C genomic DNA content are derived from the symbiotic nodule cells. We calculated an ER index from the flow cytometry measurements to express the relative amounts of 16C, 32C, and 64C nuclei in nodules (Maunoury et al. 2010) and compared it among the two *M. truncatula* accessions infected with the four rhizobia strains. We observed a high ER index for the efficient A17 nodules infected with the FSM-MA and WSM419 strains (Fig. 4A) and R108 nodules infected with the FSM-MA and 102F34 strains (Fig. 4B), while the other, less-efficient nodules had a low ER index. The number of the highly endoreduplicated cells represented by the ER index were in agreement with the histological analysis. These results indicate a strong correlation between the formation and maintenance of endoreduplicated symbiotic cells and the efficiency of the nodules.

Finally, we also analyzed the differentiation of the bacteroids in the symbiotic cells of the symbiotic couples at 5 wpi. Bacteroids were extracted from the nodules, were stained with PI, and their DNA content was measured by flow cytometry and compared with bacterial cultures of the corresponding strains (Fig. 5A and B). In the A17 nodules, bacteroids with a high DNA content were formed when the inoculation was performed with the FSM-MA and WSM419 strains (Fig. 5A and B). In contrast, in the R108 nodules, bacteroids displayed a higher ploidy level with FSM-MA and 102F34 compared with Sm1021 and WSM419 (Fig. 5A and B). These results demonstrate a correlation between polyploidy level of the bacteroids and the symbiotic efficiency of the interactions.

Incomplete differentiation of the symbiotic cells explains suboptimal symbiotic efficiency of Sm1021 in *M. truncatula* subsp. *tricycla* R108 nodules.

To gain a more precise understanding of the causes underlying the differences in symbiotic efficiencies between the *Sinorhizobium* strains, we compared different macroscopic, histological, cytological, physiological, and molecular parameters in nodules. We focused on the associations of Sm1021 and FSM-MA with R108, because these two interactions were among the most contrasting in the previous analysis. Observations were made at 7, 14, and 21 days postinoculation (dpi). Already at 7 dpi, better growth could be observed for R108 plants inoculated with FSM-MA than for those inoculated with Sm1021, and this difference became more accentuated at the two later time points (Fig. 6A). This difference could not be explained by different kinetics of nodule induction, because both plant groups developed the same number of nodules per plant (Fig. 6B). However, the phenotype of the nodules was dramatically different. At all time points, nodules induced by FSM-MA were larger and more intensely colored (Fig. 6C). At 7 dpi, FSM-MA nodules were already pinkish, suggesting that they were active in nitrogen fixation, while Sm1021 nodules were still white or only very faintly pink. These macroscopic observations were corroborated by ARAs, which revealed a higher activity for the FSM-MA-infected nodules than for those infected with Sm1021 at all time points (Fig. 6D). Already at 7 dpi, nitrogenase activity was detectable in FSM-MA nodules while it was close to background levels in Sm1021 nodules. Live/Dead BacLight staining and confocal microscopy revealed that the histological organization of nodules induced by the two strains was very similar (Fig. 6E). Already at 7 dpi, nodules infected with both strains displayed typical zonation, with meristem, infection zone, and the zone with symbiotic nodule cells entirely filled with differentiated bacteroids. Nevertheless, determination of the ER index by flow cytometry showed that the level of differentiation of the symbiotic cells in the Sm1021 nodules was lower than in the FSM-MA nodules (Fig. 6F). Thus differences that cannot be observed by microscopy are readily measured by the quantitative flow cytometry analysis of the ER index. Together, these observations demonstrate that both strains induce nodules with equal efficiency, that the organogenesis of nodules infected by both strains is proceeding similarly, but that the differentiation of the symbiotic cells in Sm1021-infected nodules is not advancing to the same extent as in the FSM-MA-infected nodules. To further explore this idea, we analyzed, by reverse transcription-quantitative polymerase chain reaction (RT-qPCR), the expression of marker genes (Fig. 6G; Supplementary Fig. 5). Based on temporal expression profiles from microarray transcriptome data (Maunoury et al. 2010) and spatial profiles obtained from laser-capture microdissection coupled to RNA sequencing (Roux et al. 2014), we selected genes that are expressed during early nodule development and in the apical part of the nodule, i.e., in either the meristem, the infection zone, or both (apical markers), genes that are expressed in the infection zone and fixation zone (intermediate markers), and finally, genes that are expressed in the fixation zone only (proximal markers) (Fig. 6G). Interestingly, the apical and intermediate markers were expressed at similar levels in both types of nodule or even more strongly in the Sm1021-infected ones. However, the proximal markers (those expressed in the symbiotic cells) were expressed at much lower levels in the Sm1021-infected nodules than in the FSM-MA-infected nodules, confirming that the Sm1021-infected symbiotic cells are less differentiated than the FSM-MA-infected symbiotic cells. The level of symbiotic cell differentiation is correlated with a strongly reduced expression of the nitrogenase-encoding *nifH* gene in the Sm1021 bacteroids (Fig. 6G), explaining the suboptimal efficiency of the Sm1021 nodules.

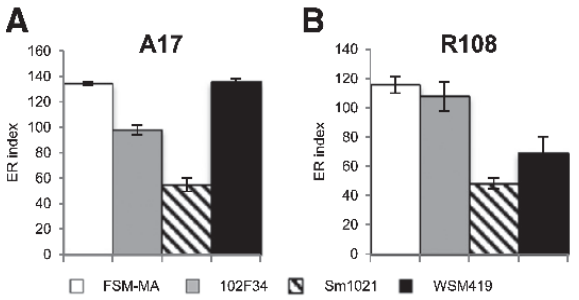


Fig. 4. Nodule ploidy analysis of *Medicago truncatula* A17 and R108 inoculated with different *Sinorhizobium* strains. **A,** Endoreduplication (ER) index of nodule nuclei derived from *M. truncatula* A17. **B,** ER index of nodule nuclei derived from *M. truncatula* R108. Plants were at 3 weeks postinoculation and were inoculated with the different rhizobia strains FSM-MA, 102F34, Sm1021, and WSM419. The ER index expresses the relative amounts of 16C, 32C, and 64C nuclei in nodules.

Senescence as a mechanism to control the interactions between *M. truncatula* accessions and *Sinorhizobium* strains.

To investigate whether the differential efficiency of the *M. truncatula* host–*Sinorhizobium* strain combinations is also correlated with nodule senescence, we characterized the senescence status of 7-wpi nodules of the A17 and R108 plants infected with the four *Sinorhizobium* strains (Supplementary Fig. 6). A17 nodules induced by FSM-MA or WSM419 were big and pink, indicating that those nodules were healthy, whereas those infected with 102F34 or Sm1021 were small and light green, indicative for an advanced stage of senescence. The green color originates from the accumulation of biliverdin, a degradation product of the heme cofactor of leghemoglobin. R108 formed elongated nodules with all the

Sinorhizobium strains. However, pink nodules were observed on R108 plants inoculated with FSM-MA or 102F34, whereas the nodules infected with WSM419 or Sm1021 were green at their base.

To further assess a potential link with senescence at a molecular level, we analyzed the expression of the senescence markers *MtCP2* and *MtCP3* (Van de Velde et al. 2006) in the 7-wpi nodules of *M. truncatula* A17 and R108 accessions infected with the different *Sinorhizobium* strains (Fig. 7A and B). Low expression of *MtCP2* and *MtCP3* was observed in nodules infected with FSM-MA on both plants, in A17 nodules infected with WSM419, and R108 nodules infected with 102F34,

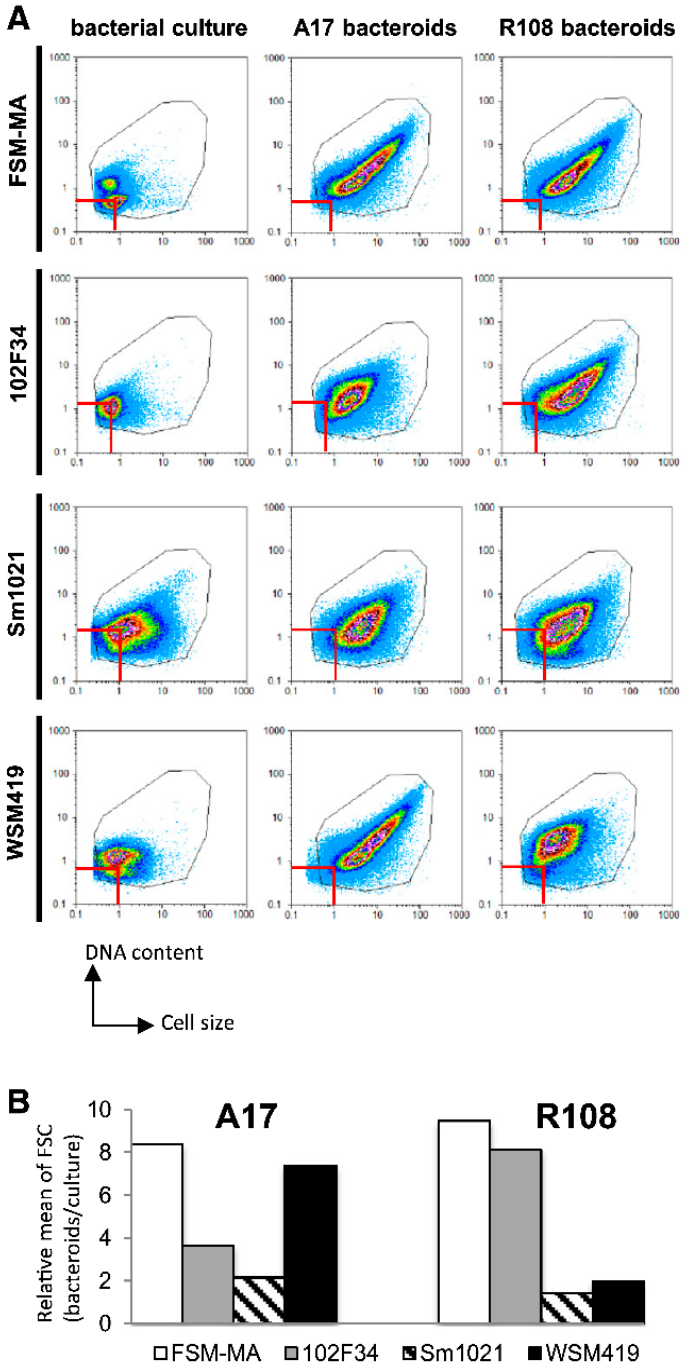


Fig. 5. Bacteroid differentiation analysis of *Medicago truncatula* A17 and R108 inoculated with different *Sinorhizobium* strains. **A,** Flow cytometry analysis of propidium iodide (PI)-stained bacterial cultures of strains FSM-MA, 102F34, Sm1021, and WSM419 or their respective bacteroids isolated from nodules formed on *M. truncatula* A17 and R108. The *x* axis represents the forward scatter (cell size) and the *y* axis corresponds to PI fluorescence (DNA content). The red lines indicate the mean position of the cultured bacteria and highlight the cell enlargement and DNA increase in the bacteroids. **B,** Increase in cell size of bacteroids, compared with cultured bacteria. The ratio of the mean forward scatter (FSC) of bacteroids to the mean FSC of the cultured bacteria is represented.

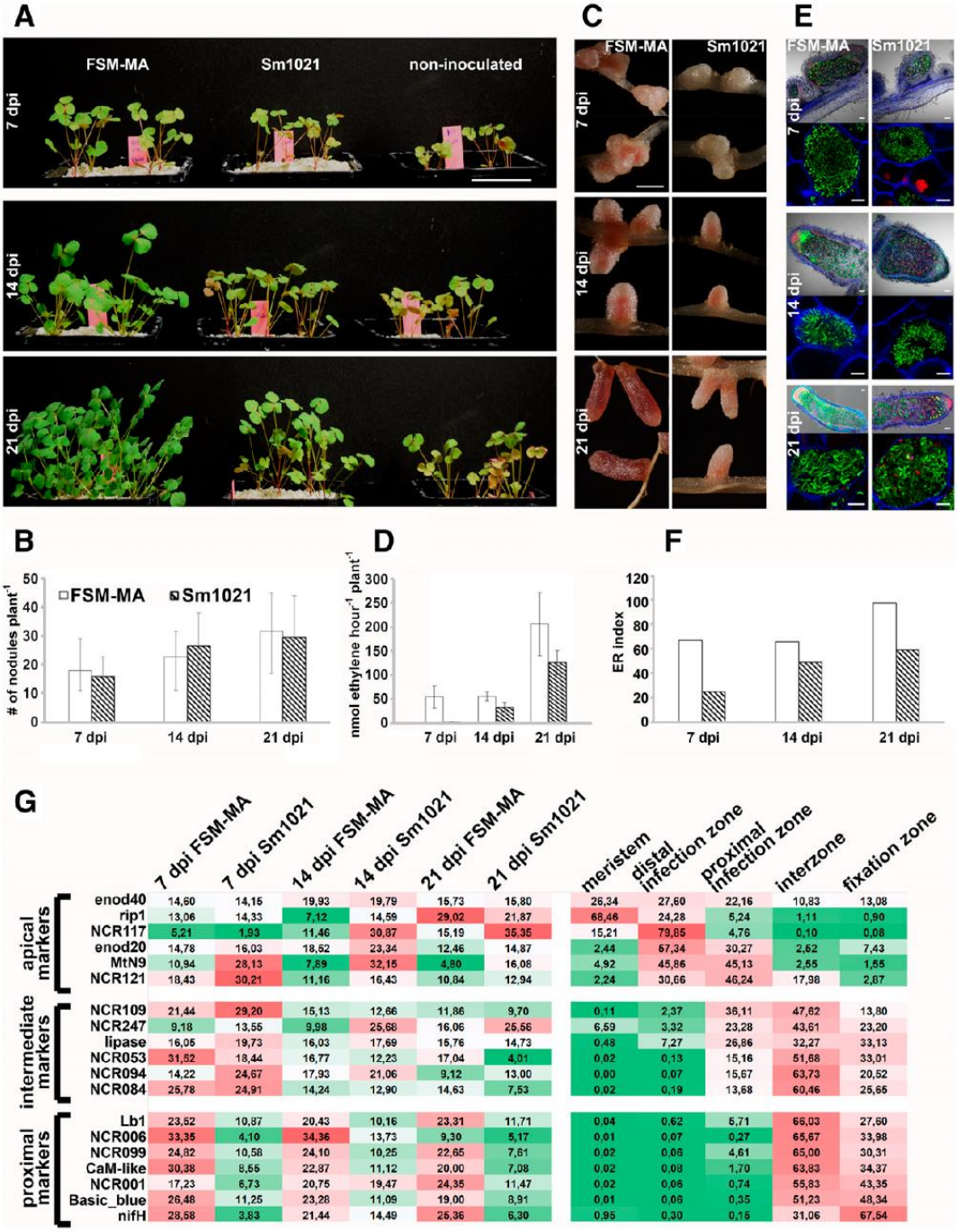


Fig. 6. Kinetics of nodule development in FSM-MA and Sm1021 infected *Medicago truncatula* subsp. *tricycla* R108. **A,** Phenotypes of *M. truncatula* R108 at 7, 14, and 21 days postinoculation (dpi) with FSM-MA or Sm1021. All the plants were photographed at the same scale. The scale bar represents 5 cm. **B,** Number of nodules formed per plant on *M. truncatula* R108 infected by FSM-MA or Sm1021 at 7, 14, and 21 dpi. Values represent the mean ± standard deviation (n = 50). **C,** Nodule phenotypes from *M. truncatula* R108 inoculated with FSM-MA or Sm1021. **D,** Total nitrogenase activity (nmol ethylene per hour per plant) of *M. truncatula* R108 infected by FSM-MA or Sm1021 and at 7, 14, and 21 dpi. Values represent the mean ± standard error (n = 10). **E,** Histological organization of nodules from *M. truncatula* R108 inoculated with FSM-MA or Sm1021. Nodule sections were stained with the Live/Dead BacLight probe marking viable bacteroids with green fluorescence from SYTO9 and dead bacteria or plant nuclei in red fluorescence by propidium iodide. Representative confocal microscopy pictures of whole nodules or of details of infected cells in the zone III of nodules are shown. The scale bars represent 100 µm (whole nodules) or 10 µm (infected cells). **F,** Nodule ploidy analysis of *M. truncatula* R108 inoculated with FSM-MA or Sm1021 at 7, 14, and 21 dpi. The Endoreduplication (ER) index expresses the relative amounts of 16C, 32C, and 64C nuclei in nodules. **G,** Gene expression heat map of nodule marker genes. Green is weak, red is high relative expression and pale colors are intermediate levels. Expression levels are indicated in the cells and expressed as percentage of the total overall samples. Genes are ordered according to their spatial expression profile in nodules with the top maximum expression in the nodule apex (meristem) and at the bottom maximum expression in the nitrogen-fixation zone of the nodule (root proximal). The genes are grouped in three clusters: apical markers, intermediate markers, and proximal markers. Left heat map, expression determined by reverse transcription-quantitative polymerase chain reaction in *M. truncatula* R108 nodules, at 7, 14, and 21 dpi, infected with FSM-MA or Sm1021. Right heat map, expression determined by RNA-Seq, data extracted from Roux et al. (2014).

whereas a high accumulation of *MtCP2* and *MtCP3* transcripts was observed in nodules of both accessions infected with Sm1021, in nodules of A17 infected with 102F34, and in nodules of R108 infected with WSM419. This is in agreement with the visual observation of senescence in nodules and reveals an inverse correlation between activation of nodule senescence and symbiotic efficiency.

DISCUSSION

In this study, we have compared morphological, histological, cytological, physiological, and molecular parameters in eight different symbiotic associations, formed by two *M. truncatula* accessions and four *Sinorhizobium* strains. In summary, we find i) a strong plant genotype-bacterial strain variability in symbiotic efficiency, a feature that has been well-documented previously (Laguerre et al. 2012; Larrainzar et al. 2014; Terpolilli et al. 2008); ii) a strong correlation between the symbiotic efficiency of the association and the degree of differentiation of both the symbiotic nodule cells and the bacteroids; iii) that, at least in the case of the interaction of Sm1021 with *M. truncatula* R108, the suboptimal interaction is not the result of poor nodule initiation or nodule organogenesis but of a lower level of differentiation of the symbiotic cells; and iv) that the most widely used strain, Sm1021, performs poorly on both tested *M. truncatula* accessions but that a newly identified strain, FSM-MA, is an excellent symbiont for the two *M. truncatula* accessions as well as for the tested *M. sativa* cultivars.

Differentiation of symbiotic nodule cells and bacteroids is a major determinant of symbiotic efficiency.

It is important to note that all eight *M. truncatula–Sinorhizobium* couples investigated here produced functional, nitrogen-fixing nodules that were infected with rhizobia. The histological analysis by microscopy and flow cytometry indicated that, in all cases, nodules had the typical organization in the distinct zones and that infected symbiotic cells were formed containing differentiated bacteroids. These observations showed that nodule organogenesis is, in all cases, proceeding normally. At least in the case of infection of R108 with strains Sm1021 and FSM-MA, the kinetics of nodule formation is very similar. The most striking difference between strongly and weakly fixing nodules was the level of symbiotic cell differentiation measured by flow cytometry and expressed by the ER index of the nodules. In the case of infection of R108 with strains Sm1021 and FSM-MA, the different level of differentiation was obvious from the earliest measured time point. The lower differentiation of Sm1021-infected symbiotic cells was also obvious from gene expression analysis of marker genes that are specifically expressed in those cells. Moreover, we demonstrated that the degree of bacteroid differentiation, characterized by the level of cell enlargement and DNA amplification, is also positively correlated with the nitrogen-fixation capacity of the tested combinations. These observations suggest thus that the degree of differentiation of both the symbiotic nodule cells and the bacteroids are variable factors that have a strong impact on the efficiency of the symbiosis.

The differentiation of the bacteroids in *M. truncatula* is determined by the NCR peptides. The symbiotic cells in *M. truncatula* produce a remarkable diversity of several hundred different NCR peptides, each encoded by a distinct gene (Mergaert et al. 2003). Some of these peptides manipulate the cell cycle of the infecting bacteria and induce them into elongated, polyploid bacteroids (Penterman et al. 2014; Van de Velde et al. 2010). However, the large diversity and the fact that many of the NCR peptides continue to be produced after the bacteroid differentiation has been completed, suggest that these peptides may have additional functions in the bacteroids (Kereszt et al. 2011). Interestingly, it was recently shown that there is a high variation in the expression level and nucleotide sequence of *NCR* genes among *M. truncatula* accessions, including R108 and A17 (Nallu et al. 2014). However, the authors compared the inoculation with two different *Sinorhizobium* strains and observed only subtle differences in *NCR* expression, suggesting that the plant does not tailor *NCR* expression according to its symbiotic partner. In contrast, we find, here, that the expression of 10 tested NCR peptides substantially differs between Sm1021-infected and FSM-MA-infected nodules. Thus, it is tempting to speculate that the variations in differentiation of the four *Sinorhizobium* strains are, in part, related to the differences in NCR production by the A17 and R108 accessions, to bacterial strain-specific profiles of NCR peptides produced by the symbiotic cells, as well as to different responses of these strains to NCR peptides. Nevertheless, we cannot exclude that the differences in *NCR* expression are the consequence of the distinct differentiation processes rather than their cause.

The fact that many legume species lack *NCR* genes and, therefore, do not impose a bacteroid differentiation process by cell elongation and genome amplification demonstrates that this bacterial differentiation process is not required per se for the symbiosis to be functional. On the other hand, the enormous investment of the host plant in NCR production has been taken for an indication that the NCR-mediated bacteroid differentiation process improves the efficiency of the symbiosis (Kereszt et al. 2011; Kondorosi et al. 2013). Our observation that the efficiency of the symbiosis correlates with the degree of bacteroid differentiation is in agreement with this hypothesis. It will be of great interest to determine if the connection between the nitrogen-fixation capacity and bacteroid differentiation by elongation and DNA amplification can be generalized to other legume-rhizobium interactions that have a similar bacteroid differentiation mechanism, i.e., in legumes of the inverted repeat-lacking and Dalbergoid clades (Czernic et al. 2015; Guefrachi et al. 2015; Kondorosi et al. 2013).

A notable difference between strongly and weakly fixing nodules from our analyses was the maintenance of the symbiotic cells. In the efficient symbiotic couples, older nodules were entirely filled with bacteroid-containing symbiotic cells, contrary to the less efficient ones, in which only a few cell layers of symbiotic cells were present while the older parts of the nodule seemed to be weakly infected. Strikingly, senescence of nodule tissues is exacerbated in the less-efficient host plant–bacteria

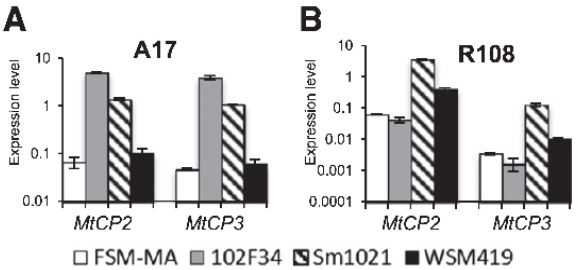


Fig. 7. Expression of the senescence-related nodulation markers *MtCP2* and *MtCP3* in nodules from *Medicago truncatula* R108 and A17 inoculated with the different *Sinorhizobium* strains at a late developmental stage. **A,** *MtCP2* and *MtCP3* expression in A17 nodules 7 weeks postinoculation (wpi) with the bacterial strains FSM-MA, 102F34, Sm1021, and WSM419, using reverse transcription-quantitative polymerase chain reaction. **B,** *MtCP2* and *MtCP3* expressions in 7-wpi R108 nodules. Bars represent standard error from three biological replicates of pools of nodules from four to eight plants and two technical replicates.

combinations. This senescence possibly corresponds to a process of "sanctions" imposed by the host on the symbionts, with incomplete bacteroid and symbiotic nodule cell differentiation. Such ineffective or less-effective nitrogen fixers can be viewed, from the plant perspective, as cheaters that attempt to unduly gain benefit from the host (Kiers et al. 2003; Oono et al. 2011). The senescence could be induced by the suboptimal nitrogen supply provided by the bacteria to its plant partner, similarly to the physiological process of senescence activation in leaves by low nitrogen supply (Balazadeh et al. 2014). Determining the level of nodule bacteria capable of surviving from senescent nodules would be necessary to demonstrate that the senescence observed in the inefficient symbiotic couples indeed corresponds to host sanctions.

Strain differences affecting symbiotic efficiency.

A previous study described strain-genotype specificity in the *S. meliloti–M. truncatula* interaction with a different set of strains and plant accessions and has attributed this specificity to structural differences of the succinoglycan oligosaccharides produced by the studied strains (Simsek et al. 2007). Succinoglycan is an exopolysaccharide produced by *S. meliloti*. Moreover, overproduction of succinoglycan has been shown to slightly improve the symbiotic productivity of *S. meliloti* Sm1021 on *M. truncatula* A17 (Jones 2012). We noticed that, when grown on agar plates, the four strains included in this study greatly differed in the production of exopolysaccharide-containing slime. Notably, strains FSM-MA and WSM419 were very strong producers while Sm1021 and 102F34 were not. Therefore, we cannot exclude a contribution of differences in succinoglycan production by *Sinorhizobium* strains in the here-described specific interactions. Nevertheless, exopolysaccharide production cannot be the only factor determining the different performance of the strains, because the weak exopolysaccharide–producing strain 102F34 performs better on *M. truncatula* R108 than the strong producer WSM419.

The different behavior of the four strains could be related to a different sensibility of the strains to the bacteroid-inducing NCR peptides. The bacterial targets for the NCR peptides and the molecular mechanism that lead to the differentiated bacteroid state remain largely unknown. Nevertheless, a few key actors have been identified. These are BacA, a peptide transporter providing protection against the antimicrobial activity of the NCR peptides (Haag et al. 2011), CtrA, a master cell cycle regulator whose inactivation by NCR peptides (by direct or indirect interactions) results in the inhibition of cell division and the amplification of the DNA content (Penterman et al. 2014; Pini et al. 2015), and peptidases, such as HrrP produced by certain rhizobial strains, which degrade NCR peptides, allowing rhizobia to escape or reduce differentiation (Price et al. 2015). Other cellular targets of NCR peptides have been identified, but their precise role in the bacteroid differentiation process is still not entirely clear (Farkas et al. 2014). It is not unlikely that molecular differences in one or more of these NCR targets, reducing or enhancing their interaction with the host NCR peptides, are at the origin of the different levels of differentiation of the four strains analyzed here.

We recently obtained the genome sequence of the novel strain FSM-MA (*unpublished*). Its sequence is very close to the sequence of Sm1021, with an average sequence identity of 99.4%, but, interestingly, the most important difference between the two strains is located in the pSymA chromid, which is about 80,000 bp larger in strain FSM-MA. The chromid pSymA is the location of the major symbiotic functions of *Sinorhizobium meliloti* strains. Comparative genomics between Sm1021 and FSM-MA and, particularly, between their pSymA chromids has the potential to identify, in FSM-MA, novel bacterial symbiotic genes and genes that are associated with the exceptional symbiotic performance of this strain.

***S. meliloti* FSM-MA as a strain of choice for the study of nodulation in *M. truncatula*.**

In conclusion, we propose the recently sequenced FSM-MA strain as a good choice for future symbiosis research with both *M. truncatula* A17 and R108. Past genetic screens have successfully identified symbiotic mutants in both *M. truncatula* A17 and R108 infected with strain Sm1021 or its sister strain Sm2011 (Pislariu et al. 2012; Starker et al. 2006). However, it is possible that mutants have been overlooked, by the use of these strains, that could have been identified using a more efficient strain such as FSM-MA. In addition, strain FSM-MA could be of great utility in reverse genetics approaches that use RNA interference-based gene knockdown to study the symbiotic role of candidate genes. The use of a highly efficient strain such as FSM-MA might be particularly critical for studies that focus on late stages of nodule development, including host cell differentiation, bacteroid differentiation, or nodule senescence, because the less-efficient strains, as we show here, induce nodules with suboptimal differentiation and early senescence, even in the wild-type host. Bacterial symbiotic mutants, which are extensively used to dissect the symbiotic process, are not yet available in the FSM-MA strain. However, this shortcoming should not be a serious obstacle in light of the available genome sequence of the strain, the ease of the creation of mutants in *Sinorhizobium* strains, and the broad knowledge of *S. meliloti* symbiotic functions that could be translated to this novel interesting strain. Finally, future genetic screens in the bacterial partner, such as screens with the novel powerful method known as transposon insertion sequencing or Tn-seq (Chao et al. 2016), aiming to identify bacterial genetic determinants for nodule occupancy, bacteroid differentiation, chronic infection, and nitrogen fixation, could be better done with a strain such as FSM-MA rather than Sm1021.

MATERIALS AND METHODS

Bacterial strains and growth conditions.

Four *Sinorhizobium* strains were used in this study: *S. meliloti* Sm1021 (Galibert et al. 2001); *S. meliloti* 102F34, a commercial strain from Nitragin (Milwaukee, WI) with unknown origin that has been used in older studies on *Medicago* nodulation (Ditta et al. 1980; Pocard et al. 1997); *S. meliloti* FSM-MA (Faculté Sciences Meknes), which was isolated from *Medicago arborea* nodules by one of us (J. Ibijbijen, *personal communication*) and was recently sequenced (*unpublished*), and *S. medicae* WSM419, which was isolated from *Medicago murex* nodules (Reeve et al. 2010).

Bacterial strains were grown in YEB medium (0.5% beef extract, 0.1% yeast extract, 0.5% peptone, 0.5% sucrose, 0.04% $MgSO_4 \cdot 7H_2O$, pH 7.5) at 30°C. Growth media were supplemented with the appropriate antibiotics, streptomycin at 100 $\mu g \cdot ml^{-1}$ for *S. meliloti* Sm1021 and chloramphenicol at 25 $\mu g \cdot ml^{-1}$ for *S. medicae* WSM419. *S. meliloti* FSM-MA and *S. meliloti* 102F34 were grown without antibiotics. For plant inoculation, cultures were centrifuged and bacterial pellets were resuspended in water at an optical density at 600 nm (OD_{600}) = 0.05.

Plant material and growth conditions.

Two accessions of *Medicago truncatula*, the Jemalong A17 reference (Young et al. 2011) and R108 (Hoffmann et al. 1997), were used in this study. The *Medicago sativa* accessions were

the commercial cultivars Nagyszenas (Hungary), Sitel (France), and Gabès (Tunisia).

For nodulation experiments, *Medicago* seeds were scarified with sulphuric acid for 3 min, were surface-sterilized with bayrochlore (Bayrol, chlorifix, 0.825% wt/vol) for 20 min, were soaked in sterile water, and were then stratified for 5 days at 4°C in the dark on solid medium (Bacto agar, 1.5%), to ensure uniform germination. Seedlings were transferred to perlite/sand (3:1 vol/vol) for a week (24°C, photoperiod 16 h of light and 8 h of dark, humidity 60%) and were inoculated with 50 ml of the appropriate bacterial culture (OD_{600} = 0.05) per pot, each pot containing five plantlets. Plants were watered with a commercial N-free fertilizer (PlantProd solution [N-P-K, 0-15-40; Fertil] at 1 g per liter).

For macroscopic and microscopic analysis, seedlings were transferred to square petri dishes (12 × 12 cm) containing solid Fahreus medium (Truchet et al. 1985). After 5 days, plantlets were inoculated with the appropriate rhizobium strain using 10 ml of bacterial suspension (OD_{600} = 0.05) per plate.

Nitrogenase activity assay.

Nitrogen-fixation activity was evaluated using the ARA as described by Koch and Evans (1966). Whole nodulated root systems were collected at the indicated times after rhizobium inoculation and were incubated with 250 µl of acetylene for 1 h at room temperature in a 10-ml air-tight glass vial with rubber stop. After incubation, 250 µl of gas mixture was removed from the vial and was injected in a gas chromatography system (7820A; Agilent Technology) to evaluate ethylene production, assessed as the area under the ethylene peak in the chromatogram. Two biological replicates were carried out with 10 plants per condition. The nitrogenase activity was calculated as ARA units (nanomoles ethylene per hour per plant and nanomoles ethylene per milligram of fresh nodule weight). Negative controls were tested by incubating 250 µl of acetylene without plant and noninoculated plants, which never produced detectable levels of ethylene.

***nifH* promoter activity.**

Three-day-old *M. truncatula* seedlings were inoculated with the four different rhizobia strains, each carrying the *nifH::GUS* reporter construct. Nodules were collected at 3 wpi and were fixed with 4% formaldehyde in 1× phosphate buffered saline (PBS) (pH 7.4) for 30 min on ice, were rinsed three times for 10 min in 1× PBS (pH 7.4), and were embedded in agarose 5% (wt/vol) (SeaKem LE agarose; Lonza). Tissue sections of 50 µm were prepared with a Zeiss MICROM HM650V vibratome. GUS staining was carried out as described by Vanstraelen et al. (2009) and was allowed to proceed for 15 min at 37°C. Nodule sections were examined under a Zeiss Axiovert 40 CFL microscope with 10× objective.

Macroscopic analysis.

Plants were photographed and collected 5 wpi for macroscopic analysis. Nodulated roots and aerial parts were weighed (fresh weight).

In vivo Live/Dead staining and confocal microscopy.

Three-week-old nodules from in vitro cultivated plants were harvested, were embedded in agarose 6% (wt/vol), and were prepared for Live/Dead staining as described by Haag et al. (2011). Nodule sections (70 µm) were prepared with a Leica VT1200S vibratome (Leica Microsystems GmbH) and were incubated for 20 min in Live/Dead staining solution (5 µM SYTO9 and 30 µM PI in 50 mM Tris-HCl buffer, pH 7.0 [Live/Dead BacLight, Invitrogen]) containing 0.01% Calcofluor White M2R (Sigma). Sections were removed from the staining solution, were rinsed of excess dye in water, and were mounted in 50 mM Tris-HCl buffer, pH 7.0, for microscopic observations. Images were taken using a Leica confocal laser scanning microscope TCS SP2.

Bacteroid isolation.

Bacteroids were purified as previously described (Mergaert et al. 2006). Nodules were harvested in ice-cold bacteroid extraction buffer (BEB) (125 mM KCl, 50 mM Na-succinate, 50 mM *N*-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid, pH 7.0) containing 0.1% bovine serum albumin and was ground with a precooled mortar and pestle. The slurry was centrifuged at 100 × *g* for 10 min to remove plant debris. The upper phase was centrifuged at 3,600 × *g* for 10 min, the supernatant was removed, and the bacteroid pellet was resuspended in BEB. The bacteroids were further purified by Percoll density gradient centrifugation. The bacteroid suspension (100 to 500 µl) was layered on the top of 7 ml of 70% Percoll in an ultracentrifuge tube and was centrifuged for 30 min at 48,000 × *g* at room temperature. The bacterial band was carefully removed from the upper part of the gradient, using a Pasteur pipette, was diluted in excess BEB, and was centrifuged again at 3,600 × *g* for 10 min. Supernatant was removed and the bacteroid pellet was resuspended in BEB.

Flow cytometry measurements.

Nodule ploidy levels were assessed as described (Maunoury et al. 2010). Nodules were chopped with a razor blade in buffer (45 mM $MgCl_2$, 30 mM trisodium citrate, 20 mM 3-(*N*-morpholino)propanesulfonic acid, 0.1% Triton X-100, pH 7.2 to 7.4) to break cells and release nuclei. Suspensions were then filtered through 30-µm Celltrics filters (Partec). Nuclei were stained by 4′,6-diamidino-2-phenylindole (DAPI) (5 µg ml^{-1}) and was analyzed using a Beckman MoFlow Astrios (Beckman Coulter) flow cytometer with a 488-nm laser for scattering and a 355-nm laser for DAPI excitation. Nuclei were gated in a DNA/side scatter plot (DAPI versus SSC) according to their DNA content and granularity. ER index was calculated based on the level of the 16C, 32C, and 64C endoreduplicated nodule nuclei according to the following formula: (% 16C nuclei) × 3 + (% 32C nuclei) × 4 + (% 64C nuclei) × 5 (Maunoury et al. 2010).

Prior to flow cytometry, purified bacteroids were fixed by heat (70°C for 15 min) and were stained with PI (50 µg ml^{-1}). PI fluorescence was analyzed with excitation at 561 nm, using a Partec CyFlow Space and the Summit 6.2 software. Minimally, 100,000 single events were measured per sample, and a PI fluorescence histogram was used to assess the DNA ploidy level of the bacteroids. Cultured bacteria of the four strains were used as reference samples.

RT-qPCR analysis.

Total RNA was isolated from *M. truncatula* nodules (1 mg) using the Plant RNeasy mini kit (Qiagen) and was subjected to DNase I treatment (Fermentas). The first-strand cDNA was synthesized from 1 µg of total RNA, using the SuperScript II first strand synthesis kit, according to the manufacturer's instructions (Invitrogen). First-strand cDNA of plant RNA was obtained with a poly-dT primer, while cDNA of bacterial RNA was obtained with a random hexamer primer. Real-time RT-qPCR reactions were performed on a Roche Light Cycler 480, using the Light Cycler fast start DNA master SYBR green I kit according to manufacturer's instructions (Roche). Cycling conditions were as follows: 95°C for 10 min, 50 cycles at 95°C for 5 s, 58°C for 5 s, and 72°C for 15 s. Three independent biological experiments and two technical replicates (based on two independent cDNA syntheses derived from the same RNA sample) were performed in all cases. The two used plant reference genes, *MtACT* (actin 11) and *MtH3L* (histone 3-like)

(Plet at al. 2011), were previously defined using GeNorm software. Primers used for *MtACT* expression: ACCCAAAG CATCAAATAATAAGTCAACC (forward) and TGGCATCA CTCAGTACCTTTCAACTG (reverse). Primers used for *MtH3L* expression: ATTCCAAAGGCGGCTGCATA (forward) and CTTTGCTTGGTGCTGTTTAGATGG (reverse). Primers used for *MtCP2* expression: CATCTTACCCTACTGCTTAAA TGC (forward) and AACTAGAAACCATGATGAATGTAGC (reverse). Primers used for *MtCP3* expression: TGGAAGC ATCTTACCCTACTG (forward) and ATATACATAAATCG CAAATCACATTC (reverse) (Pérez Guerra et al. 2010). The primers for the nodule marker genes were reported previously (Maunoury et al. 2010). The reference for bacterial gene expression was the 16S ribosomal RNA amplified with the primers GCAGAACCTTACCAGCCCTT (forward) and TATC ACCGGCAGTCCCCTTA (reverse). The *nifH* genes of Sm1021 and FSM-MA differ only in two nucleotides, and the primers CGGCGTTATCACCTCGATCA (forward) and CGATGTTG TTGGCGGCATAG (reverse) amplified a 100% identical region of the gene. Expression levels were calculated using three biological replicates and two technical replicates, normalized with the expression of the reference genes.

Statistical test.

For each experiment, two independent biological replicates with 10 to 15 plants each were performed. A nonparametric Kruskal-Wallis test (available in XLSTAT software) was used.

ACKNOWLEDGMENTS

We thank J. Ronfort and J.-M. Prosperi for providing *M. truncatula* seeds. We are indebted to M. Amar (CNRST Rabat, Morocco) and P. Ratet (IPS2, France) for providing bacterial strain FSM-MA. We thank M. Bourge (Imagerie Gif, I2BC, Gif-sur-Yvette, France) for help with the flow cytometry experiments. Imagerie Gif is supported by the "Infrastructures en Biologie Santé et Agronomie" (IBiSA), the "Agence Nationale de la Recherche" (ANR), under Investments for the Future programs "France-BioImaging infrastructure" (ANR-10-INSB-04-01), "Saclay Plant Sciences" (ANR-10-LABX-0040-SPS), and also the Conseil Général de l'Essonne. The IPS2 and I2BC institutes benefit from the support of the LabEx Saclay Plant Sciences-SPS (ANR-10-LABX-0040-SPS). T. Kazmierczak, F. Lamouche, and Q. Barrière were supported by Université Paris-Sud, and T. Kazmierczak also by the Agence Nationale de la Recherche. M. Nagymihály was supported by a fellowship from Campus France and I. Guefrachi by the France-Tunisia "Comité Mixte de Coopération Universitaire" (CMCU), project14G0817. Work by the F. Frugier team (V. Gruber and T. Kazmierczak) was supported by the Agence Nationale de la Recherche, the LabEx Saclay Plant Sciences-SPS, and the Lidex Plant Phenotyping and Engineering Pipeline-3P. The P. Mergaert team was funded by grant ANR-13-BSV7-0013.

LITERATURE CITED

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.

Balazadeh, S., Schildhauer, J., Araújo, W. L., Munné-Bosch, S., Fernie, A. R., Proost, S., Humbeck, K., and Mueller-Roeber, B. 2014. Reversal of senescence by N resupply to N-starved *Arabidopsis thaliana*: transcriptomic and metabolomic consequences. J. Exp. Bot. 65:3975-3992.

Chao, M. C., Abel, S., Davis, B. M., and Waldor, M. K. 2016. The design and analysis of transposon insertion sequencing experiments. Nat. Rev. Microbiol. 14:119-128.

Cheng, X., Wang, M., Lee, H. K., Tadege, M., Ratet, P., Udvardi, M., Mysore, K. S., and Wen, J. 2014. An efficient reverse genetics platform in the model legume *Medicago truncatula*. New Phytol. 201:1065-1076.

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., Pierre, O., Fardoux, J., Chaintreuil, C., Nguyen, P., Gressent, F., Da Silva, C., Poulain, J., Wincker, P., Rofidal, V., Hem, S., Barrière, Q., Arrighi, J. F., Mergaert, P., and Giraud, E. 2015. Convergent evolution of endosymbiont differentiation in Dalbergioid and inverted repeat-lacking clade legumes mediated by nodule-specific cysteine-rich peptides. Plant Physiol. 169:1254-1265.

Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. 1980. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. U.S.A. 77: 7347-7351.

Farkas, A., Maróti, G., Durgő, H., Györgypál, Z., Lima, R. M., Medzihradszky, K. F., Kereszt, A., Mergaert, P., and Kondorosi, É. 2014. *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. Proc. Natl. Acad. Sci. U.S.A. 111:5183-5188.

Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M.-H., Lin, Y.-H., Reid, D. E., and Gresshoff, P. M. 2010. Molecular analysis of legume nodule development and autoregulation. J. Integr. Plant Biol. 52:61-76.

Galibert, F., Finan, T. M., Long, S. R., Puhler, 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., Dreano, S., Federspiel, N. A., Fisher, R. F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R. W., Jones, T., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M. C., Pohl, T. M., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thebault, P., Vandenbol, M., Vorholter, F. J., Weidner, S., Wells, D. H., Wong, K., Yeh, K. C., and Batut, J. 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science 293:668-672.

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., and 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.

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., and Ferguson, G. P. 2011. Protection of *Sinorhizobium* against host cysteine-rich antimicrobial peptides is critical for symbiosis. PLoS Biol. 9:e1001169.

Hoffmann, B., Trinh, T. H., Leung, J., Kondorosi, A., and Kondorosi, E. 1997. A new *Medicago truncatula* line with superior in vitro regeneration, transformation, and symbiotic properties isolated through cell culture selection. Mol. Plant-Microbe Interact 10:307-315.

Jones, K. M. 2012. Increased production of the exopolysaccharide succinoglycan enhances *Sinorhizobium meliloti* 1021 symbiosis with the host plant *Medicago truncatula*. J. Bacteriol. 194:4322-4331.

Kereszt, A., Mergaert, P., and Kondorosi, E. 2011. Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims? Mol. Plant-Microbe Interact 24:1300-1309.

Kiers, E. T., Rousseau, R. A., West, S. A., and Denison, R. F. 2003. Host sanctions and the legume-rhizobium mutualism. Nature 425:78-81.

Koch, B., and Evans, H. J. 1966. Reduction of acetylene to ethylene by soybean root nodules. Plant Physiol. 41:1748-1750.

Kondorosi, E., Mergaert, P., and Kereszt, A. 2013. A paradigm for endosymbiotic life: cell differentiation of *Rhizobium* bacteria provoked by host plant factors. Annu. Rev. Microbiol. 67:611-628.

Laguerre, G., Heulin-Gotty, K., Brunel, B., Klonowska, A., Le Quéré, A., Tillard, P., Prin, Y., Cleyet-Marel, J.-C., and Lepetit, M. 2012. Local and systemic N signaling are involved in *Medicago truncatula* preference for the most efficient *Sinorhizobium* symbiotic partners. New Phytol. 195: 437-449.

Larrainzar, E., Gil-Quintana, E., Seminario, A., Arrese-Igor, C., and González, E. M. 2014. Nodule carbohydrate catabolism is enhanced in the Medicago truncatula A17-*Sinorhizobium* medicae WSM419 symbiosis. Front. Microbiol. 5:447.

Liu, J., Yang, S., Zheng, Q., and Zhu, H. 2014. Identification of a dominant gene in *Medicago truncatula* that restricts nodulation by *Sinorhizobium meliloti* strain Rm41. BMC Plant Biol. 14:167.

Łotocka, B., Kopcińska, J., and Skalniak, M. 2012. Review article: The meristem in indeterminate root nodules of Faboideae. Symbiosis 58: 63-72.

Maunoury, N., Redondo-Nieto, M., Bourcy, M., Van de Velde, W., Alunni, B., Laporte, P., Durand, P., Agier, N., Marisa, L., Vaubert, D., Delacroix, H., Duc, G., Ratet, P., Aggerbeck, L., Kondorosi, E., and Mergaert, P. 2010. Differentiation of symbiotic cells and endosymbionts in *Medicago truncatula* nodulation are coupled to two transcriptome-switches. PLoS One 5:e9519.

Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., and Kondorosi, E. 2003. A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. Plant Physiol. 132:161-173.

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A. E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., and Kondorosi, E. 2006. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. Proc. Natl. Acad. Sci. U.S.A. 103:5230-5235.

Nallu, S., Silverstein, K. A., Zhou, P., Young, N. D., and Vandenbosch, K. A. 2014. Patterns of divergence of a large family of nodule cysteine-rich peptides in accessions of *Medicago truncatula*. Plant J. 78:697-705.

Oldroyd, G. E., and Dixon, R. 2014. Biotechnological solutions to the nitrogen problem. Curr. Opin. Biotechnol. 26:19-24.

Oldroyd, G. E., and Downie, J. A. 2008. Coordinating nodule morphogenesis with rhizobial infection in legumes. Annu. Rev. Plant Biol. 59:519-546.

Oono, R., Anderson, C. G., and Denison, R. F. 2011. Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. Proc. Biol. Sci. 278:2698-2703.

Penterman, J., Abo, R. P., De Nisco, N. J., Arnold, M. F., Longhi, R., Zanda, M., and Walker, G. C. 2014. Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. Proc. Natl. Acad. Sci. U.S.A. 111:3561-3566.

Pérez Guerra, J. C., Coussens, G., De Keyser, A., De Rycke, R., De Bodt, S., Van De Velde, W., Goormachtig, S., and Holsters, M. 2010. Comparison of developmental and stress-induced nodule senescence in *Medicago truncatula*. Plant Physiol. 152:1574-1584.

Pini, F., De Nisco, N. J., Ferri, L., Penterman, J., Fioravanti, A., Brilli, M., Mengoni, A., Bazzicalupo, M., Viollier, P. H., Walker, G. C., and Biondi, E. G. 2015. Cell cycle control by the master regulator CtrA in *Sinorhizobium meliloti*. PLoS Genet. 11:e1005232.

Pislariu, C. I., Murray, J. D., Wen, J., Cosson, V., Muni, R. R., Wang, M., Benedito, V. A., Andriankaja, A., Cheng, X., Jerez, I. T., Mondy, S., Zhang, S., Taylor, M. E., Tadege, M., Ratet, P., Mysore, K. S., Chen, R., and Udvardi, M. K. 2012. A *Medicago truncatula* tobacco retrotransposon insertion mutant collection with defects in nodule development and symbiotic nitrogen fixation. Plant Physiol. 159:1686-1699.

Plet, J., Wasson, A., Ariel, F., Le Signor, C., Baker, D., Mathesius, U., Crespi, M., and Frugier, F. 2011. MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in *Medicago truncatula*. Plant J. 65:622-633.

Pocard, J. A., Vincent, N., Boncompagni, E., Smith, L. T., Poggi, M. C., and Le Rudulier, D. 1997. Molecular characterization of the *bet* genes encoding glycine betaine synthesis in *Sinorhizobium meliloti* 102F34. Microbiology 143:1369-1379.

Price, P. A., Tanner, H. R., Dillon, B. A., Shabab, M., Walker, G. C., and Griffitts, 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.

Reeve, W., Chain, P., O'Hara, G., Ardley, J., Nandesena, K., Bräu, L., Tiwari, R., Malfatti, S., Kiss, H., Lapidus, A., Copeland, A., Nolan, M., Land, M., Hauser, L., Chang, Y. J., Ivanova, N., Mavromatis, K., Markowitz, V., Kyrpides, N., Gollagher, M., Yates, R., Dilworth, M., and Howieson, J. 2010. Complete genome sequence of the *Medicago* microsymbiont *Ensifer* (*Sinorhizobium*) *medicae* strain WSM419. Stand. Genomic Sci. 2:77-86.

Rose, R. J. 2008. *Medicago truncatula* as a model for understanding plant interactions with other organisms, plant development and stress biology: past, present and future. Funct. Plant Biol. 35:253-264.

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., and Gamas, P. 2014. An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. Plant J. 77:817-837.

Simsek, S., Ojanen-Reuhs, T., Stephens, S. B., and Reuhs, B. L. 2007. Strain-ecotype specificity in *Sinorhizobium meliloti–Medicago truncatula* symbiosis is correlated to succinoglycan oligosaccharide structure. J. Bacteriol. 189:7733-7740.

Starker, C. G., Parra-Colmenares, A. L., Smith, L., Mitra, R. M., and Long, S. R. 2006. Nitrogen fixation mutants of *Medicago truncatula* fail to support plant and bacterial symbiotic gene expression. Plant Physiol. 140:671-680.

Terpolilli, J. J., O'Hara, G. W., Tiwari, R. P., Dilworth, M. J., and Howieson, J. G. 2008. The model legume *Medicago truncatula* A17 is poorly matched for N_2 fixation with the sequenced microsymbiont *Sinorhizobium meliloti* 1021. New Phytol. 179:62-66.

Truchet, G., Debellé, F., Vasse, J., Terzaghi, B., Garnerone, A. M., Rosenberg, C., Batut, J., Maillet, F., and Dénarié, J. 1985. Identification of a *Rhizobium meliloti* pSym2011 region controlling the host specificity of root hair curling and nodulation. J. Bacteriol. 164:1200-1210.

Van de Velde, W., Guerra, J. C., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., Mergaert, P., Kondorosi, E., Holsters, M., and Goormachtig, S. 2006. Aging in legume symbiosis. A molecular view on nodule senescence in *Medicago truncatula*. Plant Physiol. 141:711-720.

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., and Mergaert, P. 2010. Plant peptides govern terminal differentiation of bacteria in symbiosis. Science 327:1122-1126.

Vanstraelen, M., Baloban, M., Da Ines, O., Cultrone, A., Lammens, T., Boudolf, V., Brown, S. C., De Veylder, L., Mergaert, P., and Kondorosi, E. 2009. APC/C-CCS52A complexes control meristem maintenance in the *Arabidopsis* root. Proc. Natl. Acad. Sci. U.S.A. 106:11806-11811.

Vasse, J., de Billy, F., Camut, S., and Truchet, G. 1990. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. J. Bacteriol. 172:4295-4306.

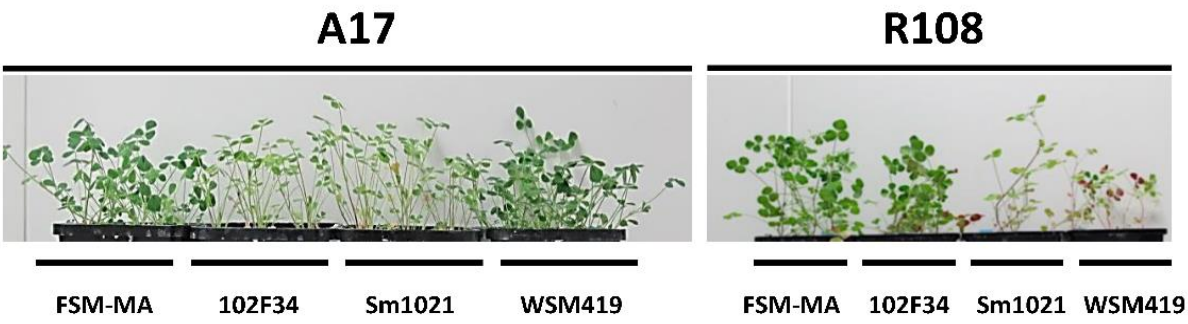
Vinardell, J. M., Fedorova, E., Cebolla, A., Kevei, Z., Horvath, G., Kelemen, Z., Tarayre, S., Roudier, F., Mergaert, P., Kondorosi, A., and Kondorosi, E. 2003. Endoreduplication mediated by the anaphase-promoting complex activator CCS52A is required for symbiotic cell differentiation in *Medicago truncatula* nodules. Plant Cell 15:2093-2105.

Yoder, J. B., Briskine, R., Mudge, J., Farmer, A., Paape, T., Steele, K., Weiblen, G. D., Bharti, A. K., Zhou, P., May, G. D., Young, N. D., and Tiffin, P. 2013. Phylogenetic signal variation in the genomes of *Medicago* (Fabaceae). Syst. Biol. 62:424-438.

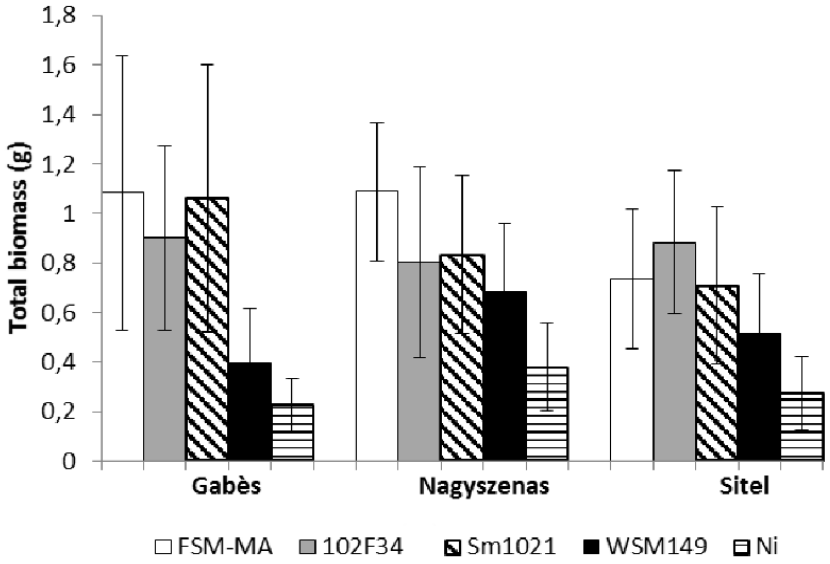
Young, N. D., Debellé, F., Oldroyd, G. E., Geurts, R., Cannon, S. B., Udvardi, M. K., Benedito, V. A., Mayer, K. F., Gouzy, J., Schoof, H., Van de Peer, Y., Proost, S., Cook, D. R., Meyers, B. C., Spannagl, M., Cheung, F., De Mita, S., Krishnakumar, V., Gundlach, H., Zhou, S., Mudge, J., Bharti, A. K., Murray, J. D., Naoumkina, M. A., Rosen, B., Silverstein, K. A., Tang, H., Rombauts, S., Zhao, P. X., Zhou, P., Barbe, V., Bardou, P., Bechner, M., Bellec, A., Berger, A., Bergès, H., Bidwell, S., Bisseling, T., Choisne, N., Couloux, A., Denny, R., Deshpande, S., Dai, X., Doyle, J. J., Dudez, A. M., Farmer, A. D., Fouteau, S., Franken, C., Gibelin, C., Gish, J., Goldstein, S., González, A. J., Green, P. J., Hallab, A., Hartog, M., Hua, A., Humphray, S. J., Jeong, D. H., Jing, Y., Jöcker, A., Kenton, S. M., Kim, D. J., Klee, K., Lai, H., Lang, C., Lin, S., Macmil, S. L., Magdelenat, G., Matthews, L., McCorrison, J., Monaghan, E. L., Mun, J. H., Najar, F. Z., Nicholson, C., Noirot, C., O'Bleness, M., Paule, C. R., Poulain, J., Prion, F., Qin, B., Qu, C., Retzel, E. F., Riddle, C., Sallet, E., Samain, S., Samson, N., Sanders, I., Saurat, O., Scarpelli, C., Schiex, T., Segurens, B., Severin, A. J., Sherrier, D. J., Shi, R., Sims, S., Singer, S. R., Sinharoy, S., Sterck, L., Viollet, A., Wang, B. B., Wang, K., Wang, M., Wang, X., Warfsmann, J., Weissenbach, J., White, D. D., White, J. D., Wiley, G. B., Wincker, P., Xing, Y., Yang, L., Yao, Z., Ying, F., Zhai, J., Zhou, L., Zuber, A., Dénarié, J., Dixon, R. A., May, G. D., Schwartz, D. C., Rogers, J., Quétier, F., Town, C. D., and Roe, B. A. 2011. The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480:520-524.

AUTHOR-RECOMMENDED INTERNET RESOURCES

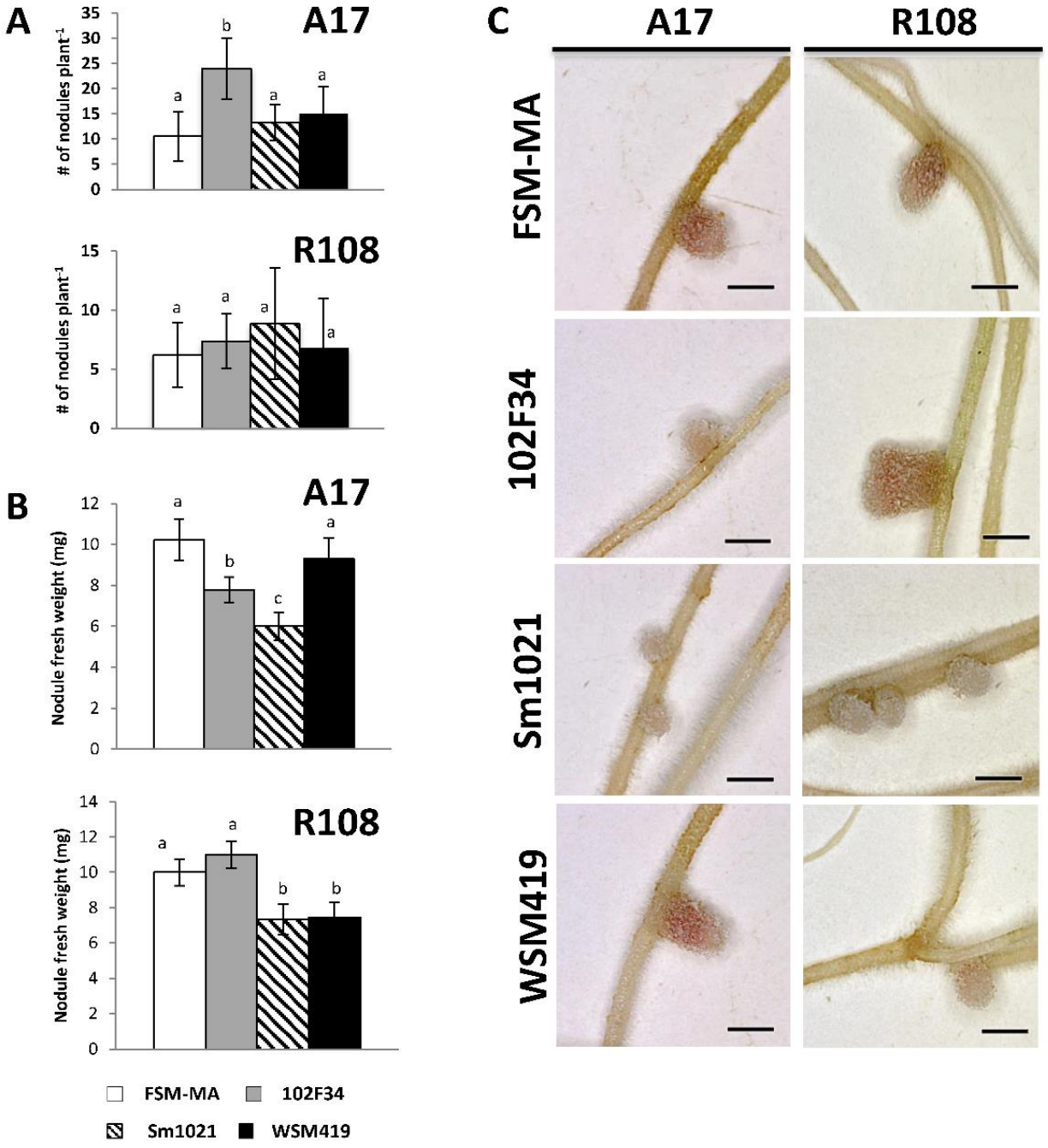
GeNorm software: http://genorm.cmgg.be
The *Medicago truncatula* Hapmap Project database: http://www.medicagohapmap.org
Xlstat software: http://www.xlstat.com/



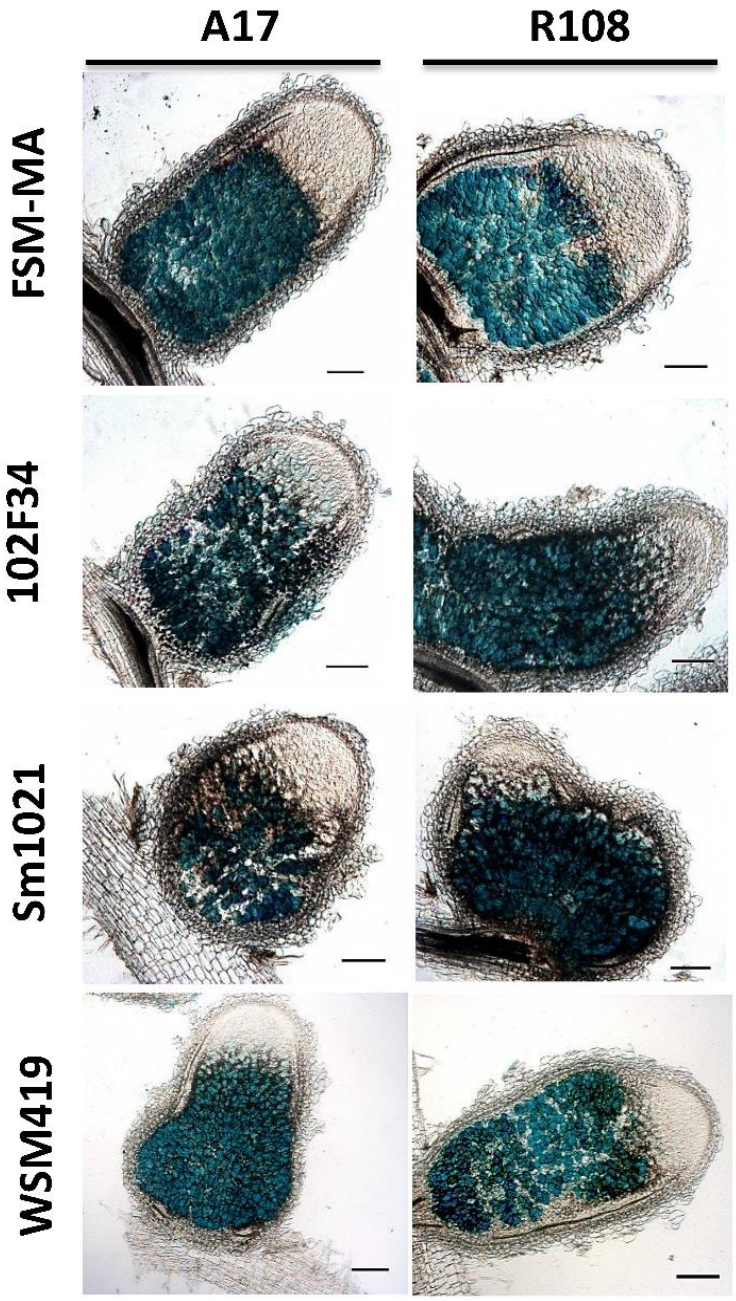
Supplemental Fig. 1. Plant growth of *M. truncatula* A17 and R108 inoculated with different *Sinorhizobium* strains. *M. truncatula* A17 and R108 plants were photographed at 5 wpi with the bacterial strains FSM-MA, 102F34, Sm1021 and WSM419.



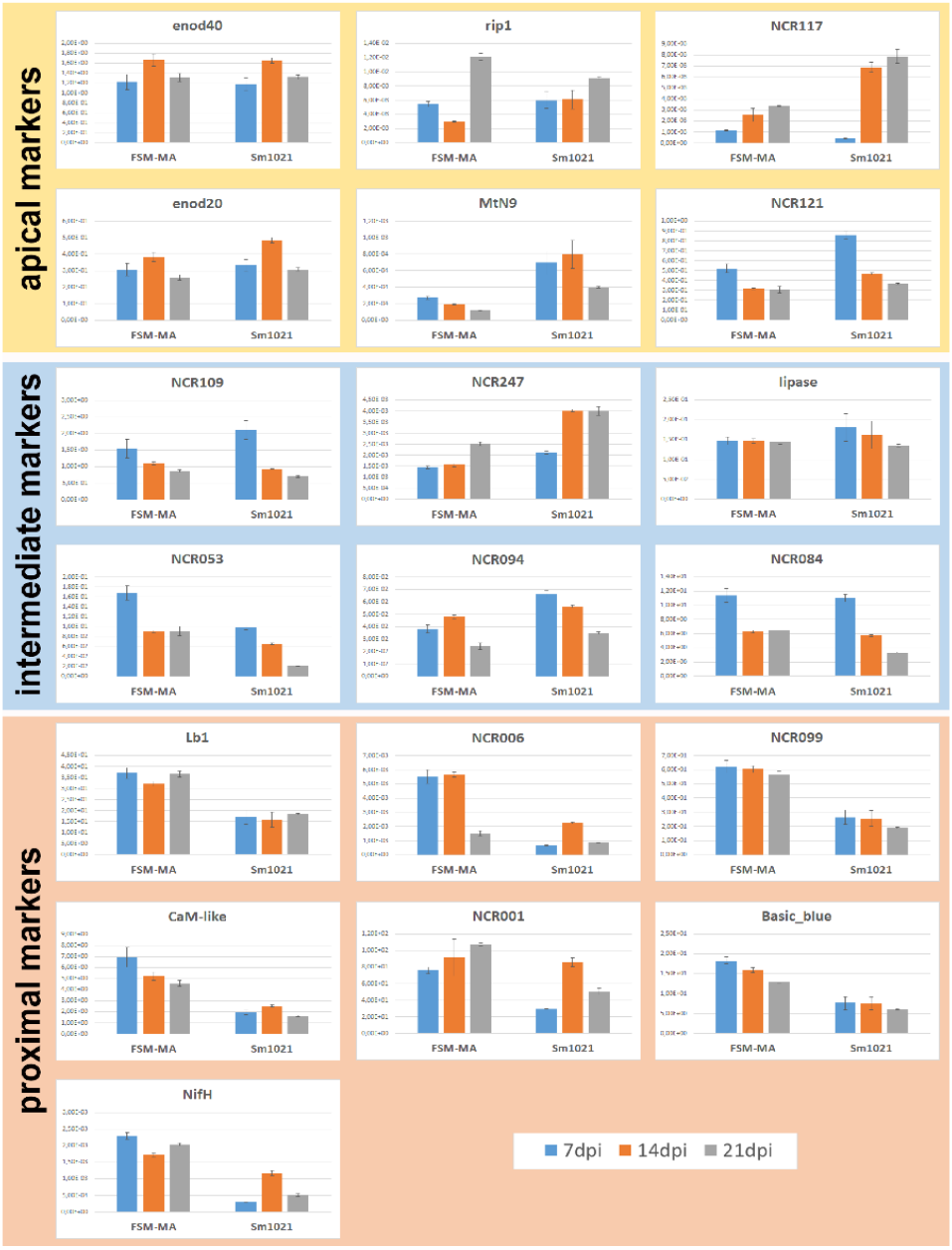
Supplemental Fig. 2. Plant biomass production of *Medicago sativa* cultivars inoculated with different *Sinorhizobium* strains. Total biomass production (fresh weight) of *Medicago sativa* cultivars Gabès, Nagyszenas and Sitel, in response to nodulation with strains Sm1021, FSM-MA, 102F34 and WSM149 or on uninoculated plants (Ni). Fifteen plants were tested per condition. Error bars indicate standard deviations.



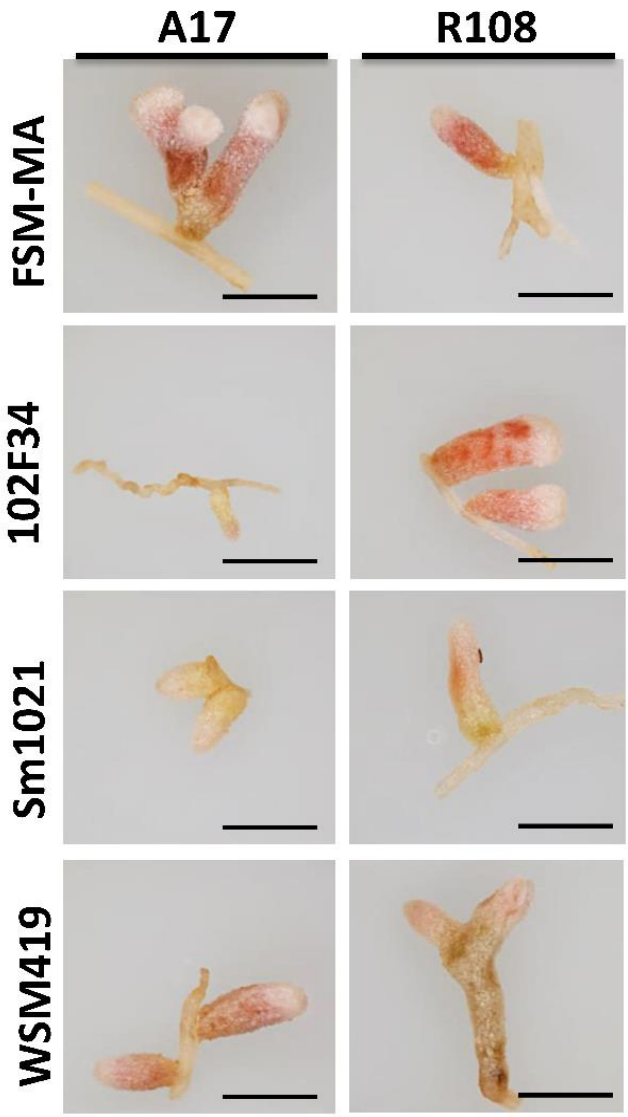
Supplemental Fig. 3. Nodule phenotypes from the *M. truncatula* A17 and R108 inoculated with different *Sinorhizobium* strains. **A,** Number of nodules per plant at 14 dpi. **B,** Total nodule biomass (fresh weight) per plant; Values represent the mean +/- standard error (n=10). Letters indicate statistical differences between groups ($p<0,05$; Kruskal-Wallis test). **C,** Representative images of nodules at 3 wpi from A17 and R108 infected with the strains FSM-MA, 102F34, Sm1021 and WSM419. The scale bar represents 500µm.



Supplemental Fig. 4. *nifH* expression in nodules from *M. truncatula* A17 and R108 inoculated with different *Sinorhizobium* strains. Representative images of the microscopic analysis of the *nifH::GUS* expression in nodules at 3 wpi from A17 and R108 infected with the strains FSM-MA, 102F34, Sm1021 and WSM419. The scale bars represent 100µm.



Supplemental Fig. 5. RT-qPCR expression analysis of marker genes in *M. truncatula* R108 nodules at 7, 14 and 21 dpi, infected with FSM-MA or Sm1021. The diagrams show the expression of the indicated genes after normalization. Bars represent standard deviation from three biological replicates of pools of nodules from 20 to 40 plants, and two technical replicates. Genes are grouped as in Fig. 6G.



Supplemental Fig. 6. Late stage of nodules from *M. truncatula* A17 and R108 inoculated with different *Sinorhizobium* strains. Representative images of nodules at 7 wpi from A17 and R108 infected with the strains FSM-MA, 102F34, Sm1021 and WSM419. The scale bar represents 500µm. Development of senescence appears as a green zone in the basal part of the nodules. Images are representative nodules obtained from four to eight plants.

2. Integrated roles of BclA and DD- carboxypeptidase 1 in *Bradyrhizobium* differentiation within NCR-producing and NCR-lacking root nodules

Quentin Barrière, Ibtissem Guefrachi, Djamel Gully, Florian Lamouche, Olivier Pierre, Joël Fardoux, Clémence Chaintreuil, Benoît Alunni, Tatiana Timchenko, Eric Giraud et Peter Mergaert

Publié dans la revue *Scientific Reports* (2017). DOI: https://doi.org/10.1038/s41598-017-08830-0

www.nature.com/scientificreports

OPEN

Received: 18 May 2017
Accepted: 12 July 2017
Published online: 22 August 2017

Integrated roles of BclA and DD-carboxypeptidase 1 in *Bradyrhizobium* differentiation within NCR-producing and NCR-lacking root nodules

Quentin Barrière[1], Ibtissem Guefrachi[1,2,4], Djamel Gully[3], Florian Lamouche[1], Olivier Pierre[1,5], Joël Fardoux[3], Clémence Chaintreuil[3], Benoît Alunni[1], Tatiana Timchenko[1], Eric Giraud[3] & Peter Mergaert[1]

Legumes harbor in their symbiotic nodule organs nitrogen fixing rhizobium bacteria called bacteroids. Some legumes produce Nodule-specific Cysteine-Rich (NCR) peptides in the nodule cells to control the intracellular bacterial population. NCR peptides have antimicrobial activity and drive bacteroids toward terminal differentiation. Other legumes do not produce NCR peptides and their bacteroids are not differentiated. Bradyrhizobia, infecting NCR-producing *Aeschynomene* plants, require the peptide uptake transporter BclA to cope with the NCR peptides as well as a specific peptidoglycan-modifying DD-carboxypeptidase, DD-CPase1. We show that *Bradyrhizobium diazoefficiens* strain USDA110 forms undifferentiated bacteroids in NCR-lacking soybean nodules. Unexpectedly, in *Aeschynomene afraspera* nodules the nitrogen fixing USDA110 bacteroids are hardly differentiated despite the fact that this host produces NCR peptides, suggesting that USDA110 is insensitive to the host peptide effectors and that nitrogen fixation can be uncoupled from differentiation. In agreement with the absence of bacteroid differentiation, USDA110 does not require its *bclA* gene for nitrogen fixing symbiosis with these two host plants. Furthermore, we show that the BclA and DD-CPase1 act independently in the NCR-induced morphological differentiation of bacteroids. Our results suggest that BclA is required to protect the rhizobia against the NCR stress but not to induce the terminal differentiation pathway.

Nodules on the roots of legume plants - and in certain cases also on the stems - are symbiotic organs that house within their cells thousands of nitrogen-fixing rhizobium bacteria, called bacteroids[1,2]. This intimate co-habitation of eukaryotic cells with bacteria requires adaptations of both the host and the micro-symbiont to maintain a cooperative state to avoid one partner succumbing to the aggression or immune responses of the other. The production of effector peptides, called Nodule-specific Cysteine-Rich (NCR) peptides by the Inverted Repeat Lacking Clade (IRLC) and Dalbergoid legumes is such a host adaptation that helps control the endosymbiotic bacterial population[3–5]. These peptides, related to innate immunity antimicrobial peptides, target the bacteria and induce them into a terminally differentiated state which is characterized by a pronounced cell enlargement resulting in either elongated or spherical bacteroids, depending on the host species. This bacterial cell enlargement is accompanied by a strong polyploidization of the bacterial genome.

[1]Institute for Integrative Biology of the Cell, UMR9198, CNRS, Université Paris-Sud, CEA, Gif-sur-Yvette, France. [2]Research Unit Biodiversity & Valorization of Arid Areas Bioressources (BVBAA), Faculty of Sciences, Gabès University, Erriadh-Zrig, 6072, Gabès, Tunisia. [3]Laboratoire des Symbioses Tropicales et Méditerranéennes, Institut pour la Recherche et le Développement, UMR IRD/SupAgro/INRA/UM2/CIRAD, Campus International de Baillarguet, TA A-82/J, 34398, Montpellier Cedex 5, France. [4]Present address: Université de Pau et des Pays de l'Adour, Pau, France. [5]Present address: Institut Sophia AgroBiotech, Sophia-Antipolis, France. Quentin Barrière and Ibtissem Guefrachi contributed equally to this work. Correspondence and requests for materials should be addressed to P.M. (email: peter.mergaert@i2bc.paris-saclay.fr)

Some NCR peptides, mostly cationic ones, have the capacity to kill rhizobia, as well as other bacteria, in agreement with their similarity to antimicrobial peptides such as defensins. Similarly to defensins, the NCR peptides affect bacterial envelope permeability which may lead to bacterial death[3]. In addition to their effects on the bacterial membranes, NCR peptides also have diverse intracellular targets[6], including the cell cycle machinery, which leads to the bacteroid genome amplification[7].

NCR-induced bacterial elongation and polyploidization and bacterial survival in the face of NCR toxicity require a SbmA_BacA transmembrane domain-containing peptide importer protein which is able to transport the NCR peptides within the bacterial cytosol. It was proposed that NCR peptide uptake reduces their toxicity by preventing their deleterious effects on the bacterial membrane[8,9]. Depending on the rhizobium species, two types of NCR transporters have been identified. In *Sinorhizobium meliloti*, the symbiont of the IRLC legumes *Medicago sativa* and *Medicago truncatula*, the NCR transporter, called BacA, consists solely of the SbmA_BacA transmembrane domain and is highly homologous to *E. coli* SbmA which drives peptide transport across the inner membrane by the membrane electrochemical gradient[10]. On the other hand, *Bradyrhizobium* spp. nodulating *Aeschynomene* spp. (Dalbergoids) use for NCR uptake the SbmA_BacA transmembrane domain-containing ABC transporter BclA which has, unlike BacA, a C-terminal ATPase domain and drives transport by ATP hydrolysis[9]. The SbmA_BacA transmembrane domain transporters are wide-spread in bacteria and are not symbiosis-specific proteins. Their bacterial household function remains poorly defined but in *E. coli* or *S. meliloti* they can import a diversity of peptides or peptide derivatives, including antimicrobial peptides that have intracellular targets[11–22].

Many legumes do not produce NCR peptides and the bacteroids in those legumes do not enlarge, do not become polyploid and are not terminally differentiated[23,24]. This is the case for legumes of the Robinioids and Millettioid clades such as *Lotus japonicus*, *Phaseolus vulgaris*, *Lablab purpureus*, *Tephrosia vogelii*, *Vigna unguiculata* and soybean (*Glycine max*). Consequently, the *bacA* gene of their symbionts, *Mesorhizobium loti*, *Rhizobium leguminosarum* bv. *phaseoli*, *Rhizobium etli* and *Sinorhizobium fredii*, respectively, can be inactivated without consequences for the symbiosis[1,8,25–28]. Based on these case studies, it was proposed that the *bacA* gene in rhizobia has only a symbiotic role when the host plant produces NCR peptides[8]. However, the role of the bradyrhizobial *bclA* gene in interactions with host plants that do not produce NCR peptides has not been studied.

Besides the SbmA_BacA transporters BacA or BclA, a peptidoglycan-modifying enzyme, a specific DD-carboxypeptidase (DD-CPase) enzyme with a peptidoglycan-binding SPOR domain and encoded by the *DD-CPase1* gene, was recently identified in bradyrhizobia as a second rhizobial factor involved in terminal bacteroid differentiation in legumes producing NCR peptides[29]. DD-CPases are periplasmic enzymes that regulate the degree of reticulation in the peptidoglycan layer and thereby affect the physical strength of the bacterial wall. In the NCR-producing legumes *Aeschynomene indica* and *Aeschynomene afraspera*, the bacteroids of the *DD-CPase1* mutant are malformed and hypertrophied. However, in the NCR-lacking soybean, the mutation does not affect the bacteroids[29].

In this study, we investigated whether *Bradyrhizobium diazoefficiens* (previously *Bradyrhizobium japonicum*) strain USDA110 requires its *bclA* gene in symbiosis with NCR-lacking and NCR-producing legume hosts. We also analyzed how the *bclA* and *DD-CPase1* determinants of terminal bacteroid differentiation interfere with each other, by creating and analyzing double mutants in the *bclA* and *DD-CPase1* genes in *B. diazoefficiens* strain USDA110 and *Bradyrhizobium* sp. strain ORS285.

Results

The *Bradyrhizobium diazoefficiens* USDA110 blr7537-encoded protein is a BclA homologue. Similar to other analyzed *Bradyrhizobium* strains[9], no *bacA* homologous gene was identified in the genome of *B. diazoefficiens* USDA110. However, the *B. diazoefficiens* gene blr7537 was identified as a homolog of the *bclA* genes of *Bradyrhizobium* spp. ORS285 and ORS278[9]. The encoded protein has 70–72% identity and 80–81% similarity to the BclA proteins of strains ORS285 and ORS278. The protein has the same structure with an N-terminal SbmA-BacA transmembrane domain and a C-terminal ATPase domain. It thus encodes a potentially functional ABC transporter. Moreover, the genes in the 3 *Bradyrhizobium* species are located in syntenic regions that extend to over 200 kb. Therefore we designated blr7537 as *bclA*. Similarly as in strains ORS285 and ORS278, the *B. diazoefficiens* locus lacks genes that potentially encode additional components of the ABC transporter, such as a periplasmic binding protein for substrate binding and delivery to the transporter.

***Bradyrhizobium diazoefficiens* USDA110 BclA is an NCR peptide transporter and functional in symbiosis.** To test the *in vitro* and *in vivo* activity of the *B. diazoefficiens* USDA110 BclA protein, a deletion mutant of the *bclA* gene was constructed. The gene was also cloned into the plasmids pMG103 and pRF771, downstream of the *trp* promoter, and introduced into the *bclA*, *bacA* and *sbmA* mutants of *Bradyrhizobium* sp. ORS285, *S. meliloti* strain Sm1021 and *E. coli* strain BW25113, respectively.

Bleomycin and Bac7 are antimicrobial compounds which have intracellular targets, DNA and the ribosomes respectively, and they require active transport to be taken up in the bacterial cells. In *E. coli*, *S. meliloti* and *Bradyrhizobium* spp. ORS285 and ORS278, the uptake is mediated by the SbmA/BacA/BclA transporters[9,17,30]. Thus strains expressing one of these transporters display a significantly increased sensitivity to bleomycin or Bac7. We find that wild type *B. diazoefficiens* strain USDA110 or the *E. coli sbmA* and *S. meliloti bacA* mutants expressing USDA110 *bclA* from the pRF771 plasmid are more sensitive to bleomycin or Bac7 than the corresponding strains lacking *bclA* (Fig. 1), in agreement with BclA of USDA110 being able to transport these peptides.

Contrary to bleomycin and Bac7, the sensitivity to antimicrobial NCR peptides is reduced in the presence of *sbmA*, *bacA* or *bclA*[8,9]. This opposite response is likely because the toxicity of the NCR peptides resides in their potential to provoke membrane permeability and loss of membrane potential, rather than in the inhibition of some intracellular process. Also the *bclA* gene of USDA110 is able to reduce the sensitivity to NCR peptides in the *S. meliloti bacA* mutant (Fig. 2a). In addition, similarly as shown before for the *S. meliloti bacA* or *Bradyrhizobium*

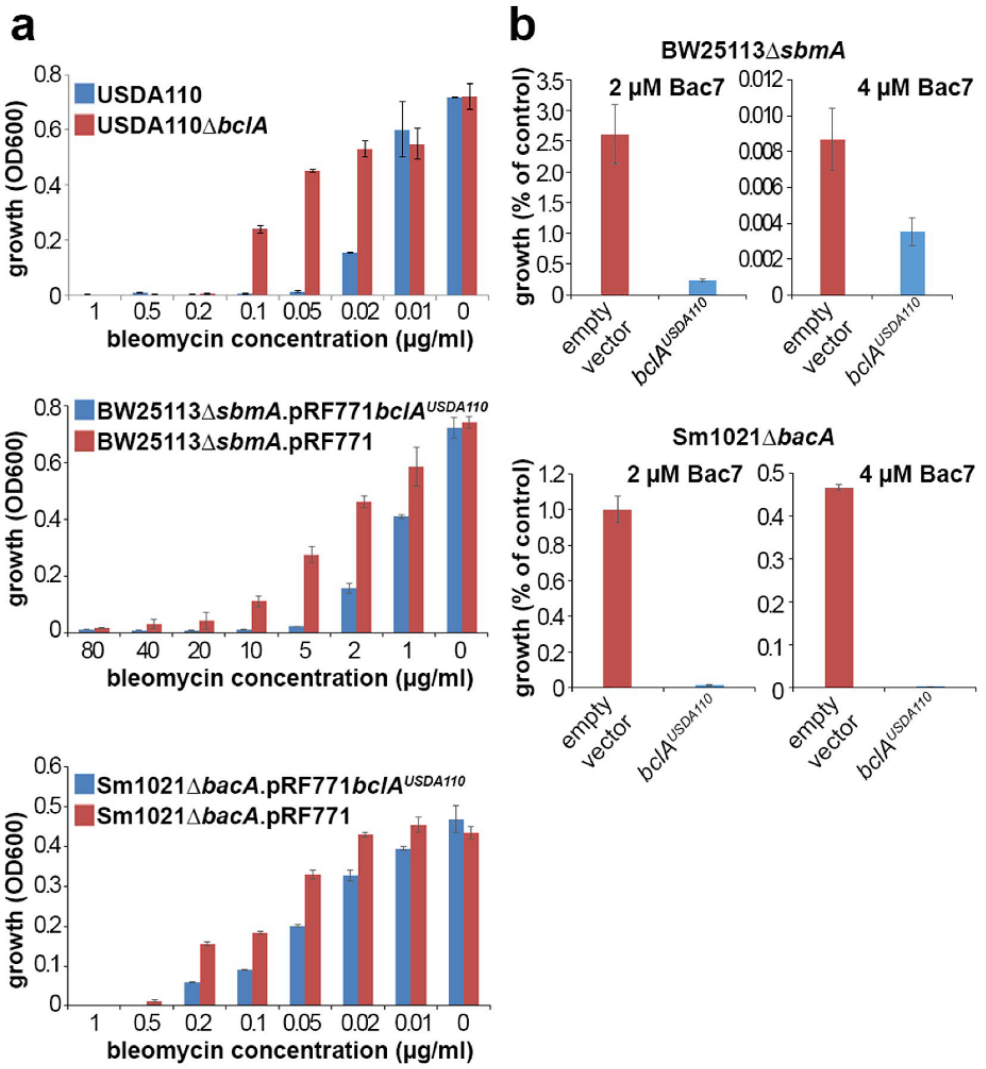


Figure 1. The USDA110 *bclA* gene confers sensitivity to the antibiotic bleomycin and the antimicrobial peptide Bac7. (**a**) Bleomycin sensitivity in *B. diazoefficiens* strain USDA110 and its *bclA* mutant derivative and in *E. coli* strain BW25113Δ*sbmA* and *S. meliloti* strain Sm1021Δ*bacA* carrying an empty vector or a vector expressing USDA110 *bclA*. Blue bars are for strains expressing *bclA* and red bars are for strains lacking the gene. Bleomycin concentrations were applied as indicated and growth was determined by optical density measurement at 600 nm with a plate reader after 72 h, 48 h or 24 h incubation for USDA110, Sm1021 and BW25113 respectively. (**b**) Survival of *E. coli* strain BW25113Δ*sbmA* (top) and *S. meliloti* strain Sm1021Δ*bacA* (bottom) derivatives carrying the empty pRF771 plasmid or the USDA110 *bclA* gene located on plasmid pRF771 after treatment with the peptide Bac7 at the indicated concentration. The surviving bacteria were counted and expressed as % from the water control treatment. Error bars in all panels are standard deviations.

strain ORS285 *bclA* genes[9], the expression of USDA110 *bclA* promotes the uptake of an FITC-modified NCR247 peptide into the *S. meliloti* strain Sm1021Δ*bacA* (Fig. 2b), while FITC alone is not taken up (Fig. 2c).

Finally, the *bclA* gene of USDA110 can complement the ORS285Δ*bclA* mutant for bacteroid differentiation in *A. indica* nodules (Fig. 3a–f). This mutant induces small nodules in which bacteria do not differentiate and die as revealed by live/dead staining of nodule sections[9] (Fig. 3c,d). The USDA110 *bclA* gene, similarly as the ORS285 *bclA* gene, and introduced into this mutant on the pMG103 plasmid, restores the wild type phenotype with the formation of large nodules inhabited with well-formed spherical bacteroids that remain alive (Fig. 3a,b,e,f). Plants inoculated with both complemented strains showed vigorous growth, contrary to those inoculated with the un-complemented mutant indicating that nitrogen fixation defect of the mutant is restored by the USDA110 *bclA* gene. Moreover, the USDA110 *bclA* gene also complements in part the Sm1021Δ*bacA* strain for nodulation of *M. sativa* (Fig. 3g–l). The nodules induced by the complemented strain become elongated and pinkish, compared to the small white nodules induced by the *bacA* mutant (Fig. 3g,i,k). The *bacA* mutant bacteria die as

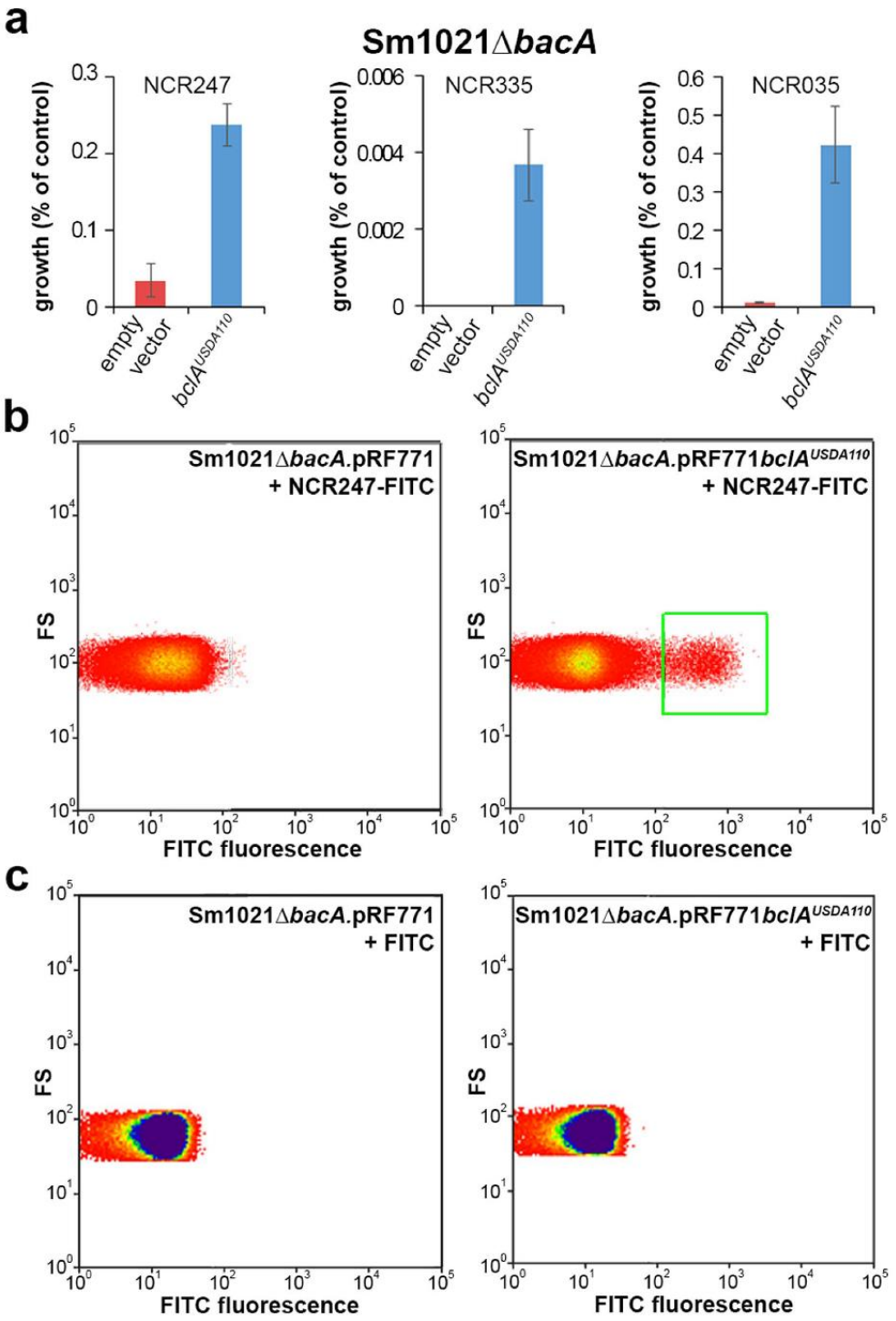


Figure 2. The USDA110 *bclA* gene confers resistance to antimicrobial NCR peptides and mediates the uptake of FITC-NCR247. (**a**) *S. meliloti* strain Sm1021Δ*bacA* derivatives expressing no *bacA*-related gene (empty vector) or the USDA110 *bclA* gene were incubated with NCR247, NCR335 or NCR035 or with water (control) and the surviving bacteria were counted and expressed as % from the control treatment. Error bars in all panels are standard deviations. (**b**) FITC-NCR247 uptake by *S. meliloti* strain Sm1021Δ*bacA* or its derivative expressing the USDA110 *bclA* gene was measured by flow cytometry in the presence of trypan blue to quench extracellular fluorescence. FITC-positive cells are marked with a green box. (**c**) FITC is not taken up by *S. meliloti* derivatives. FITC-treated bacteria were analyzed by flow cytometry in the presence of trypan blue to quench extracellular fluorescence. No FITC-positive cells were detected.

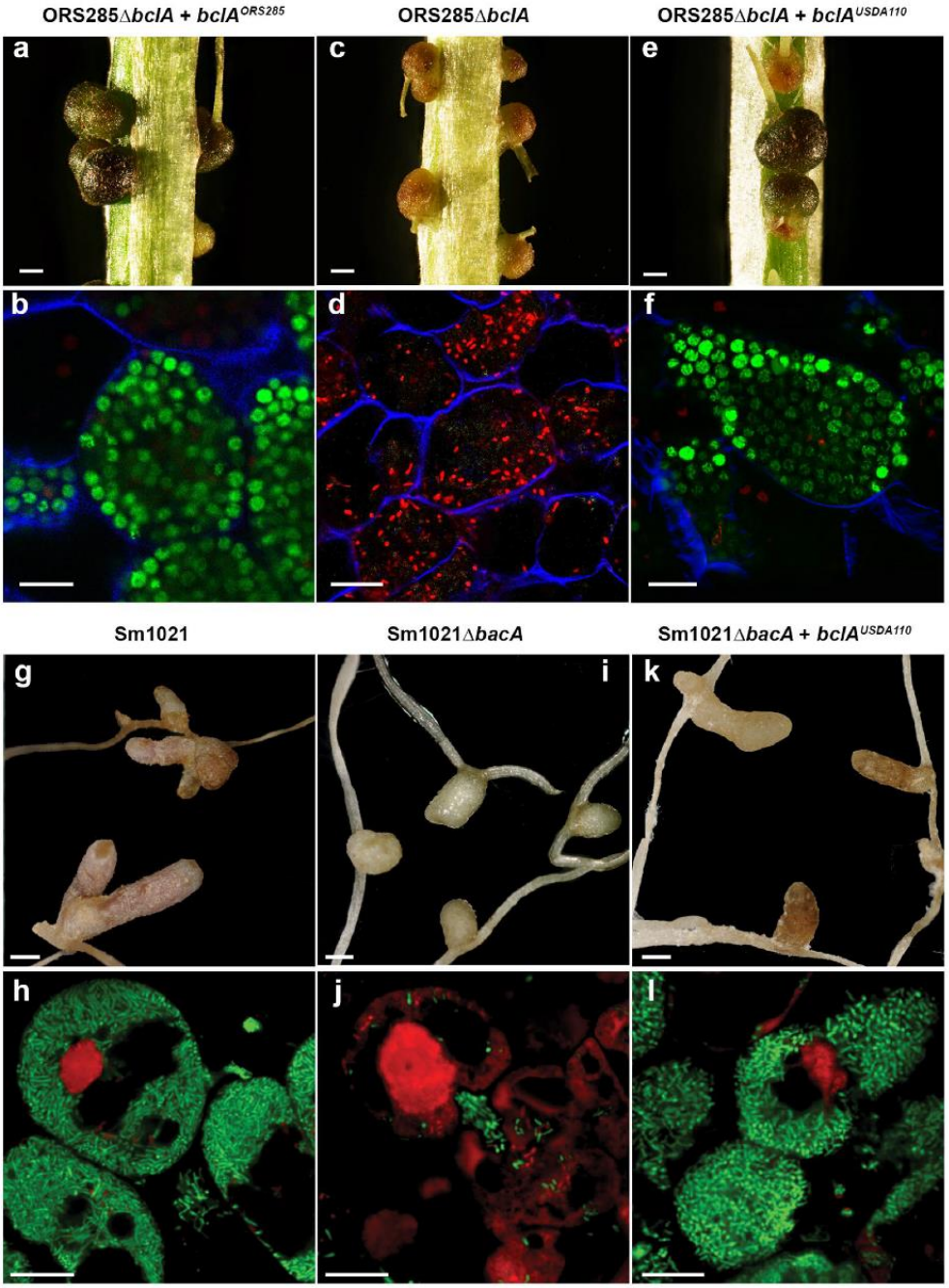


Figure 3. The *bclA* gene of *B. diazoefficiens* USDA110 complements the *bclA* mutation in *Bradyrhizobium* sp. ORS285Δ*bclA* and the *bacA* mutation in *S. meliloti* Sm1021Δ*bacA*. (**a**,**c**,**e**) Phenotype of *A. indica* nodules at 14 dpi, infected with the indicated strains. Scale bars are 1 mm. (**b**,**d**,**f**) Bacteroid viability determined by live/dead and calcofluor white staining of nodule sections and confocal microscopy in *A. indica* nodules induced by the indicated strains. Scale bars are 10 µm. (**g**,**i**,**k**) Phenotype of *M. sativa* nodules at 28 dpi, infected with the indicated strains. Scale bars are 1 mm. (**h**,**j**,**l**) Bacteroid viability determined by live/dead staining of nodule sections and confocal microscopy in *M. sativa* nodules induced by the indicated strains. Scale bars are 10 µm.

soon as they are released inside the nodule cells and exposed to the NCR peptides[8](Fig. 3j). The bacteroids of the complemented strain however are viable within the symbiotic cells indicating that the hypersensitivity of the *bacA* mutant to the NCR peptides is suppressed by the *bclA* gene of USDA110 (Fig. 3l). Nevertheless, the defect of the *S. meliloti bacA* mutation is only partially repaired by *bclA* of USDA110 because nodules do not fix nitrogen

and do not support plant growth (data not shown). This phenotype is similar to the one obtained with the *S. meliloti bacA* complementation by *bclA* of strain ORS285[9] or even by more similar *bacA* genes of *Sinorhizobium* or *Rhizobium* species[31]. This suggests that although rhizobial *bacA* and *bclA* genes have overlapping specificity for peptide uptake, they may also display differences in the set and/or amount of NCR peptides they can handle, probably because they evolved in the context of specific interactions with host plants, each producing its specific arsenal of NCR peptides. Nevertheless, the USDA110 *bclA* gene seems to be capable to treat the *Aeschynomene* NCR peptides.

Together, the bleomycin, Bac7 and NCR peptide sensitivity assays as well as the *in vivo* complementation of the *Bradyrhizobium* sp. ORS285 *bclA* mutantion or the *S. meliloti bacA* mutation indicate that the *bclA* gene of *B. diazoefficiens* strain USDA110 is functional and has a similar activity as the *bclA* gene of *Bradyrhizobium* strain ORS285 and the *S. meliloti bacA* gene.

Bradyrhizobium diazoefficiens USDA110 BclA is not required for symbiosis with NCR-lacking soybean. In agreement with the taxonomic position of soybean within the Millettioids, the *B. diazoefficiens* strain USDA110 bacteroids in soybean nodules are undifferentiated and display no cell enlargement and polyploidy as revealed by microscopy of nodule sections (Fig. 4c) and flow cytometry analysis of purified bacteroids (Fig. 4e). We tested whether the *bclA* gene in strain USDA110 is required for symbiosis with soybean. The *bclA* mutant of USDA110 was undistinguishable from the wild type strain for all parameters analyzed, including nodule tissue structure and bacterial occupation, bacterial viability, morphology, size and DNA content as well as nitrogen fixation (Fig. 4a–e). Thus the *Bradyrhizobium bclA* gene, similarly to *bacA* in other rhizobium species, is not required for symbiosis when bacteroids are not constrained by the host plant-produced NCR peptides to differentiate into an elongated and polyploid morphotype.

Bradyrhizobium diazoefficiens USDA110 BclA is also not required for symbiosis with NCR-producing _Aeschynomene afraspera_. *B. diazoefficiens* strain USDA110, which is a natural soybean symbiont, can also form functional nodules on *A. afraspera*[29,32]. The strain forms nodules composed of a central zone with fully infected symbiotic cells and a cortex layer surrounding the infected cells. This nodule organization is very similar to the histology of nodules induced by the natural *Aeschynomene* symbiont *Bradyrhizobium* strain ORS285 (Fig. S1). Nevertheless, USDA110 is a less efficient nitrogen fixer, supporting lower plant biomass production and nitrogen accumulation than ORS285[32] (Fig. S2). The latter strain forms elongated and polyploid bacteroids on *A. afraspera* and requires the *bclA* gene for elongated bacteroid formation[9]. Therefore, we analyzed the bacteroid type formed by strain USDA110 within *A. afraspera* nodules and the role of the USDA110 *bclA* gene. Unexpectedly, observations by confocal microscopy showed that USDA110 bacteroids in *A. afraspera* nodules were not or only very slightly elongated (Fig. 5b), contrary to bacteroids of strain ORS285 which are strongly elongated[4,9]. To confirm this unpredicted observation we used flow cytometry analysis of the bacteroid cell size, determined by the forward scatter (a measure for cell size), and the DNA content, measured by DAPI fluorescence. A slight increase in size of the bacteroids compared to the bacteria in culture was measured but this was not accompanied with an increase in the DNA content of the bacteroids (Fig. 5c). Thus the USDA110 bacteroids are much less or hardly differentiated compared to ORS285 bacteroids which have, besides the strong cell enlargement, also a marked increase in DNA content[4,9]. The absence of a pronounced differentiation of USDA110 bacteroids is not likely resulting from a defect in *NCR* gene expression in USDA110-infected nodules since five tested *NCR* genes were expressed at similar or even higher levels in USDA110-infected nodules compared to ORS285-infected nodules (Fig. S3).

In agreement with the absence of differentiation of the wild type USDA110, the USDA110Δ*bclA* mutant was not affected in symbiosis with *A. afraspera*: nodules infected with wild type or mutant were similar, supported plant growth and fixed nitrogen to the same extent (Fig. 5a,d), both types of nodules contained symbiotic cells which were completely infected with bacteroids that seemed not or only slightly elongated (Fig. 5b; Fig. S1) which was confirmed by flow cytometry (Fig. 5c).

BclA is not required for the formation of differentiated bacteroids in _Aeschynomene afraspera_ nodules by the USDA110 _DD-CPase1_ mutant. Contrary to the wild type USDA110 strain, the UDSA110 *DD-CPase1* mutant forms strongly elongated and polyploid bacteroids in the *A. afraspera* nodules[29] (Fig. 5b,c), indicating that by affecting their cell wall strength, the bacteria become sensitive to the NCR differentiation signals produced by the nodule cells. Nevertheless, the nitrogenase activity of plants infected with the UDSA110 *DD-CPase1* mutant is strongly reduced, in large part because the mutant induces much less nodules than the wild type[29] (Fig. 5d). We created a *bclA*/*DD-CPAse1* double mutant to determine whether the cell wall-determined bacteroid differentiation is depending on the BclA peptide transporter. Unexpectedly, the double mutant exhibited a similar symbiotic phenotype than the *DD-CPase1* single mutant (Fig. 5b–d). This result indicates that the differentiation of USDA110, made possible by the inactivation of the *DD-CPase1* gene, is independent of BclA.

A possible explanation could be that the *DD-CPase1* mutation increases the permeability of cells for peptides, including NCR peptides, rendering the BclA transporter superfluous. To test this possibility, we measured sensitivity of strains against the peptide bleomycin which needs to be internalized to target the bacterial DNA. We found that the *DD-CPase1* mutant strain displays a sensitivity to bleomycin similar to the wild type strain (Fig. 5e), indicating that the peptidoglycan remodeling, induced by the mutation, does not interfere with peptide uptake. Similarly, the double mutant is just as much or even slightly more resistant to bleomycin than the *bclA* mutant (Fig. 5e).

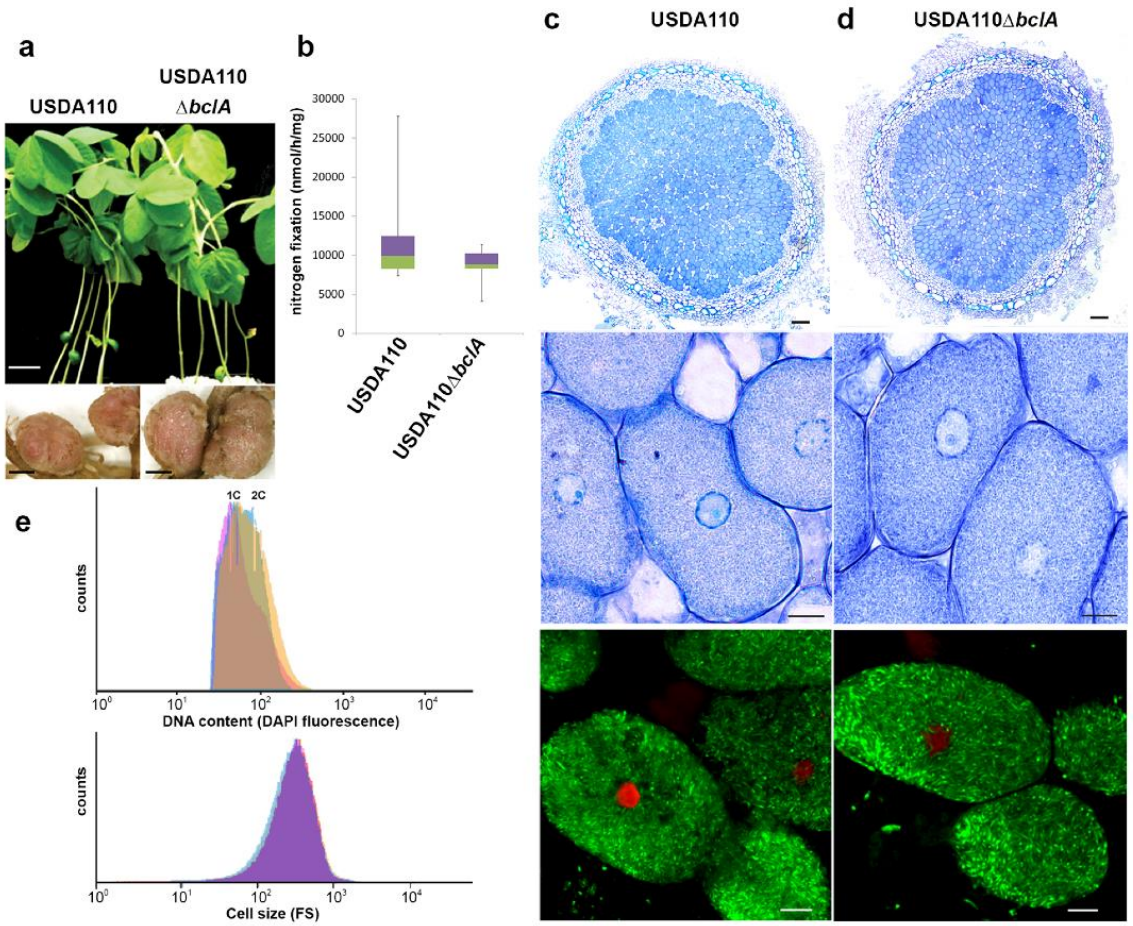


Figure 4. The *B. diazoefficiens bclA* mutant is not affected in symbiosis with soybean. (**a**) Plant growth and nodule phenotype of *G. max* plants inoculated with *B. diazoefficiens* USDA110 wild type and the *bclA* mutant at 14 dpi. Scale bars are 3.5 cm (top panels) and 1 mm (bottom panels). (**b**) The nitrogen fixation activity of *G. max* plants inoculated with *B. diazoefficiens* USDA110 wild type and the *bclA* mutant, measured by the acetylene reduction assay per mg of fresh nodule weight (nmol ethylene produced per hour incubation and per mg nodule weight). Box-plots represent in the rectangle the first quartile to the third quartile, divided by the median value, whiskers above and below the box show the minimum and maximum measured values. (**c,d**) Toluidine blue stained thin sections (two top rows) and confocal microscopy of bacteroid viability determination by live/dead staining (bottom row) of *G. max* nodules induced by USDA110 or USDA110Δ*bclA*. Scale bars are 100 µm for the top panels and 10 µm for the other panels. (**e**) Flow cytometry analysis of DNA content by DAPI fluorescence (top) and of cell size by forward scatter (bottom) in free-living *B. diazoefficiens* USDA110 (pink) or bacteroids isolated from *G. max* nodules infected with USDA110 wild type (bluc) or USDA110Δ*bclA* (orange). The forward scatter and DAPI fluorescence profiles are completely overlapping for the 3 samples.

BclA and DD-CPase1 in *Bradyrhizobium* strain ORS285 act independently in bacteroid differentiation. To further explore the interdependence of BclA and DD-CPase1 in bacteroid differentiation, we created the double mutant also in *Bradyrhizobium* strain ORS285. This strain forms nodules on *A. afraspera* as well as on *A. indica* in which it differentiates into either elongated polyploid or spherical polyploid bacteroids respectively[4]. A *bclA* mutation in this strain blocks the differentiation process in both hosts[9] (Fig. 6a) while a *DD-CPase* mutation induces hypertrophied bacteroids[29] (Fig. 6a). In a similar way as in USDA110, the ORS285 double mutant was still capable to induce strongly enlarged bacteroids in nodules of both *A. afraspera* and *A. indica* (Fig. 6a). In *A. indica*, the *bclA* mutation induced death of the bacteria as revealed by the red fluorescence in the live/dead staining procedure of nodule sections (Fig. 6a). Even if in the double mutant, many bacteroids were strongly enlarged (Fig. 6a), others remained undifferentiated. This may be related to bacterial death induced by the *bclA* mutation in such a way that many bacteria die before having the chance to enlarge. Thus, these results indicate that the *bclA* and *DD-CPase1* mutations have a cumulative effect and that the *bclA* function is not upstream of the bacterial differentiation provoked by the *DD-CPase1* mutation. The cumulative effect of the two mutations is also observed when measuring with the acetylene reduction assay the nitrogenase activity of *A. afraspera* nodules (Fig. 6b). The *bclA* mutation has a stronger impact on nitrogen fixation than the *DD-CPase1*

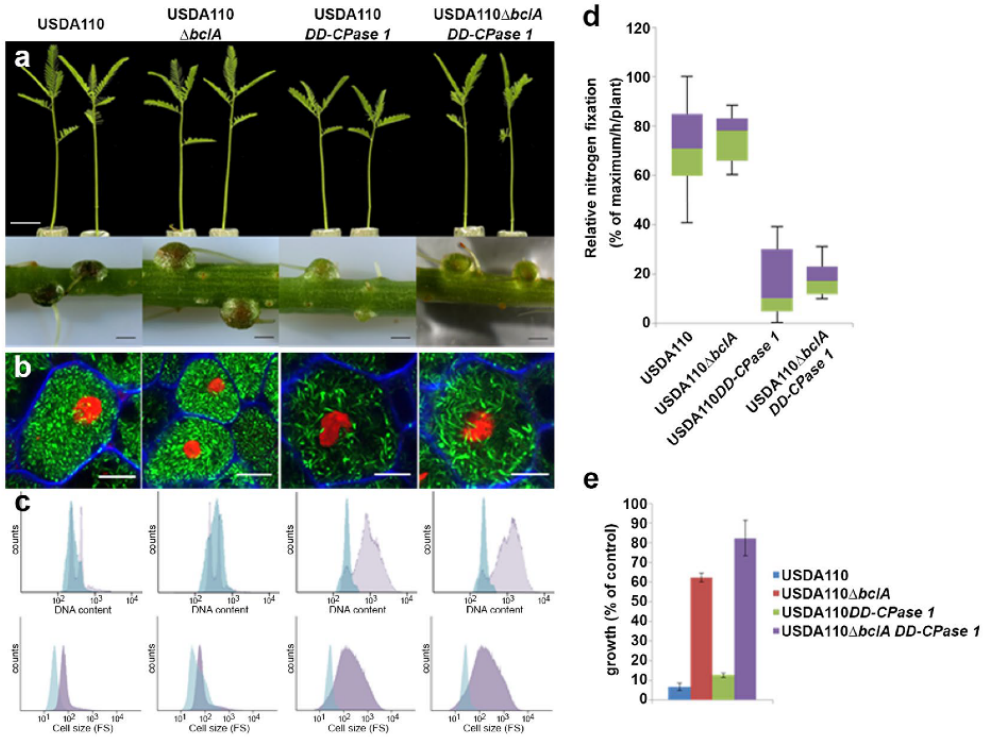


Figure 5. The *B. diazoefficiens bclA*, *DD-CPase1*, *bclA/DD-CPase1* mutant phenotypes in symbiosis with *A. afraspera*. (**a**) Plant growth and nodule phenotype of *A. afraspera* plants inoculated with *B. diazoefficiens* USDA110 wild type and indicated mutants at 14 dpi. Scale bars are 2 cm (top panels) and 1 mm (bottom panels). (**b**) Confocal microscopy of bacteroid viability determination by live/dead and calcofluor white staining of *A. afraspera* nodules induced by USDA110 wild type and indicated mutants. Scale bars are 10 µm. (**c**) Flow cytometry analysis of DNA content by DAPI fluorescence and of cell size by forward scatter (FS) in free living *B. diazoefficiens* USDA110 (blue) or bacteroids isolated from *A. afraspera* nodules infected with USDA110 wild type and indicated mutants (pink). (**d**) Relative nitrogen fixation activity (% of maximum/h/plant) of *A. afraspera* plants inoculated with *B. diazoefficiens* USDA110 wild type and indicted mutants, measured by the acetylene reduction assay per plant. Box-plots represent in the rectangle the first quartile to the third quartile, divided by the median value, whiskers above and below the box show the minimum and maximum measured values. (**e**) Bleomycin sensitivity in *B. diazoefficiens* strain USDA110 and its *bclA, DD-CPase1, bclA/DD-CPase1* mutant derivatives at 0,0375 µg bleomycin/ml. Growth was determined by optical density measurement at 600 nm with a plate reader after 72 h. Error bars in panels (d) and (e) are standard deviations.

mutation and the double mutant has the same low nitrogenase activity as the *bclA* single mutant while on the other hand bacteroids of the double mutant resemble morphologically more the *DD-CPase1* bacteroids.

Discussion

The BacA transporter is widespread but not universal in bacteria. It has been recruited by rhizobia that infect host plants producing NCR peptides in their infected nodule cells. In such host cells, the BacA protein is essential for the rhizobia to differentiate and establish a chronic infection. However, BacA is not critical for symbiosis in host cells of legumes that do not produce NCR peptides[8]. Rhizobia such as *Bradyrhizobium* spp., which lack a *bona fide bacA* gene but infect legume hosts that do produce NCR peptides in the infected nodule cells, use the different but related transporter BclA[9]. Here we show that the BclA function in *B. diazoefficiens*, in analogy with the BacA function in other rhizobia, is superfluous for chronic infection of soybean symbiotic cells that do not produce NCR peptides or when the bacterial strain is inherently non-responsive to the peptides produced by *A. afraspera* symbiotic cells. In agreement with our current findings, we previously demonstrated that the *bacA* gene of the broad host range bacterium *Sinorhizobium fredii* strain HH103 is not required for symbiosis with soybean[26]. These results, together with previously published studies support the view that the major, if not sole symbiotic role of SbmA_BacA domain transporters is the import of NCR peptides.

B. diazoefficiens strain USDA110 is not a natural symbiont of *Aeschynomene* host plants. The strain produces the same major Nod factors as the genuine *Aeschynomene* symbiont *Bradyrhizobium* strain ORS285 and is able to induce with equal efficiency nodules on *A. afraspera* roots which have a normal histological organization but a reduced nitrogen fixation and plant growth promoting activity[32] (Fig. S2). However a striking difference between the two strains is the absence of marked differentiation of the USDA110 bacteroids (this work) while the ORS285

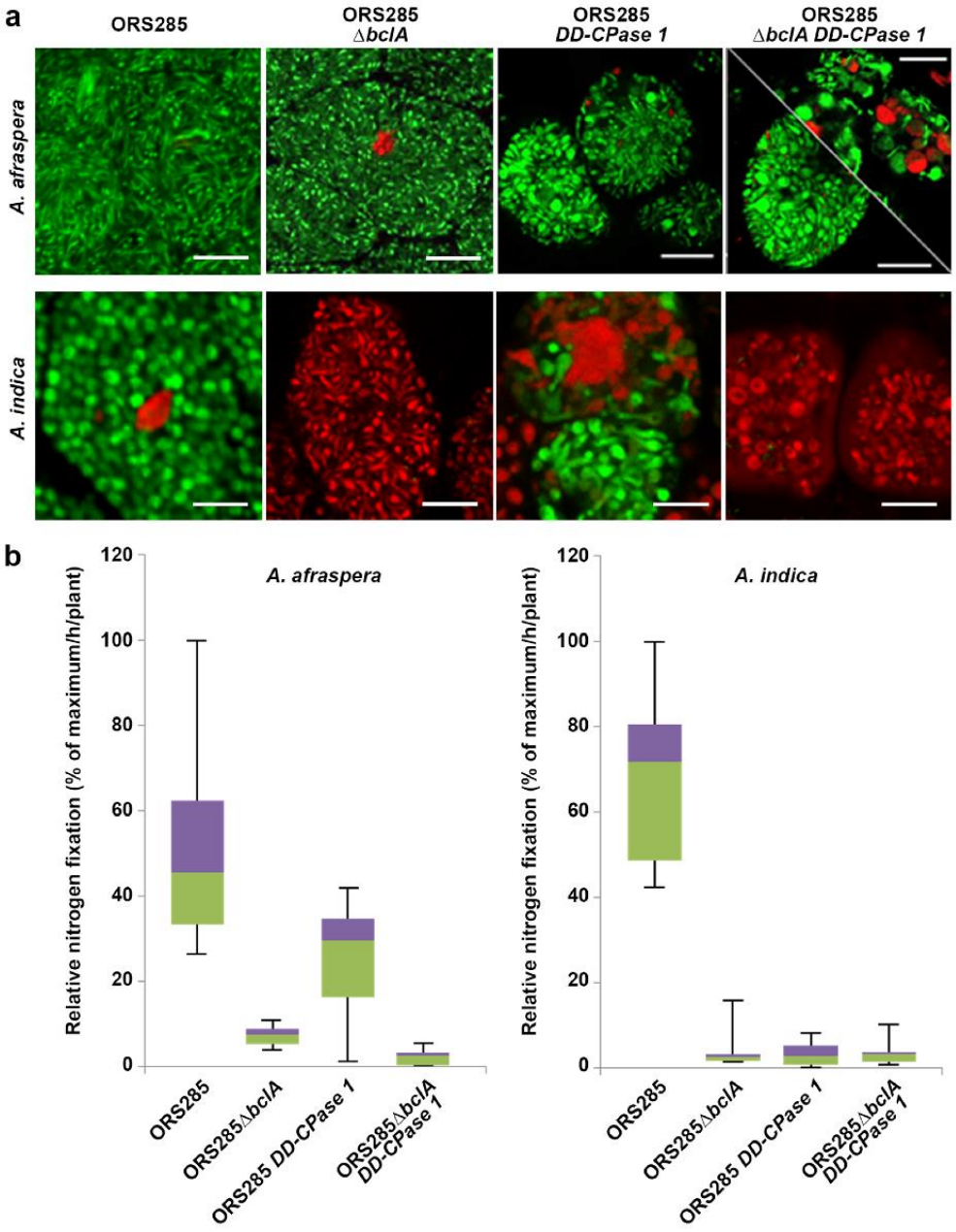


Figure 6. The *Bradyrhizobium* strain ORS285 *ΔbclA*, *DD-CPase1*, *ΔbclA/DD-CPase1* mutant phenotypes in symbiosis with *A. afraspera* and *A. indica*. (**a**) Confocal microscopy of bacteroid viability determination by live/dead staining of *A. afraspera* and *A. indica* nodules induced by ORS285 wild type and the indicated mutants. Scale bars are 10 µm. (**b**) Relative nitrogen fixation activity (% of maximum/h/plant) of *A. afraspera* and *A. indica* plants inoculated with ORS285 wild type and the indicted mutants, measured by the acetylene reduction assay per plant. Box-plots represent in the rectangle the first quartile to the third quartile, divided by the median value, whiskers above and below the box show the minimum and maximum measured values.

bacteroids strongly elongate and become polyploid[4]. The absence of differentiation of *B. diazoefficiens* strain USDA110 in *A. afraspera* nodules is unlikely to be the result of its *bclA* gene being a defective allele because we found that the USDA110 *bclA* mutant has a similar response to bleomycin as the corresponding mutants in strains ORS285 and ORS278. The USDA110 *bclA* gene fully complements the symbiotic defect of the ORS285 *bclA* mutation for bacteroid differentiation in *Aeschynomene*. It also complements *S. meliloti bacA* and *E. coli sbmA* mutants, similar to the corresponding genes of ORS285 and ORS278[9], for Bac7- and NCR-uptake and -sensitivity and for symbiosis with *M. sativa*. Together, these findings demonstrate that USDA110 BclA has the same capacity

as BclA from ORS285 or BacA from *S. meliloti* to transport a broad range of antimicrobial peptides, including Bac7, bleomycin and NCR peptides. They also tell us that USDA110 BclA does not lack the ability to uptake some specific peptide or set of peptides produced in *Aeschynomene* nodules which would explain the inability of the strain to differentiate within these nodules. Moreover, transcriptome analysis by RNA-seq (data not shown) revealed that *bclA* of USDA110 is weakly expressed, in culture as well as in soybean or *A. afraspera* nodules, but the orthologous gene of ORS285 is similarly weakly expressed. It is also unlikely that the lack of differentiation of USDA110 bacteroids results from a defect in NCR gene expression in USDA110-infected nodules since five tested NCR genes were expressed at similar or even higher levels in USDA110-infected nodules compared to ORS285-infected nodules.

A candidate factor for the strain USDA110 recalcitrance to bacteroid differentiation in *A. afraspera* nodules is the production of hopanoids[33]. Hopanoids are a bacterial class of lipids, produced by certain bacteria, notably the bradyrhizobia but not the other rhizobal genera. These lipids, which are similar to eukaryotic steroid lipids, render membranes more rigid and resistant to membrane stresses, including to antimicrobial peptides[34,35]. Strain USDA110 is remarkable among the bradyrhizobia because of its very high accumulation of hopanoids, up to 40% of the total cellular lipid fraction[33]. The amount of hopanoids accumulating in the membranes of strain ORS285 is currently unknown but it could be sufficiently low for NCR responsiveness while the high hopanoid content in USDA110 could make this strain insensitive to the NCR signals of the host. In agreement with this hypothesis, an *hpnH* mutant of USDA110, affected in the synthesis of a specific type of hopanoids, the C35 hopanoids, is drastically impaired in symbiosis with *A. afraspera* but not with soybean; the mutant, in contrast to the wild type, is also sensitive to the antimicrobial activity of NCR peptides[36]. Moreover, even if the *Bradyrhizobium* USDA110 and ORS285 strains have a high proportion of shared genes, they have nevertheless also a substantial number of unique genes and they belong to phylogenetically distinct clades in the *Bradyrhizobium* genus[37,38]. A recent comparative genome analysis of ten *Bradyrhizobium* strains, including strains USDA110 and ORS285, revealed that USDA110 has over 3000 unique genes while ORS285 had almost 500 genes not shared with the other strains included in the study[38]. Probably the different genetic repertoires of the USDA110 and ORS285 strains contribute to their differential response to the bacteroid differentiation factors produced by the symbiotic host cells. This difference could be exploited to identify new bacterial factors involved in the differentiation process and the response to NCR peptides. Whatever its molecular basis, the weak or non-responsiveness of strain USDA110 to the host-produced bacteroid differentiation factors most likely reflects the fact that this strain has co-evolved with NCR-lacking soybean and not with NCR-producing *Aeschynomene* or other NCR-producing hosts.

It is interesting to note that although the root nodules infected with the strain USDA110 fix nitrogen, they do so less efficiently when compared to nodules infected with strain ORS285; they supported less plant growth with the accumulation of lower amounts of nitrogen in the plant biomass, indicating that the undifferentiated bacteroids of USDA110 are less efficient than the differentiated bacteroids of ORS285. We recently demonstrated an analogous correlation between the level of bacteroid differentiation and symbiotic performance in the interactions of two *M. truncatula* accessions with four different *Sinorhizobium* strains[39]. These concordant observations in different legumes suggest that the differentiation of bacteroids may contribute - by a mechanism that still has to be clarified - to the nitrogen fixation efficiency of rhizobium-legume symbioses[40–42]. Variability in the response to NCR peptides can be at the basis of such variations in nitrogen fixation efficiency of *Rhizobium* strains. This has been nicely illustrated recently in the *S. meliloti*-*M. truncatula* interaction[43,44]. Allelic variations in single NCR genes between *M. truncatula* accessions can determine the nitrogen fixation effectiveness of a particular *S. meliloti* strain while not affecting other strains. Thus the genome of rhizobial symbionts should be optimally aligned with the spectrum of NCR peptides produced by the nodule cells for efficient bacteroid differentiation and nitrogen fixation.

The BclA/BacA transporters provide protection against the antimicrobial activity of NCR peptides, both *in vitro* and *in vivo*, and in addition, the transporters are also required for bacterial elongation and polyploidization[8,9] (this work). This implies two possibilities concerning the molecular function of these transporters: either their sole function is providing protection against the antimicrobial stress in the nodule cells and the lack of bacteroid differentiation in the corresponding mutants is an indirect consequence of the reduced fitness of these mutants; alternatively, the transporters are, in addition to providing protection, also directly involved in the activation of the differentiation mechanism leading to cell enlargement and polyploidization, for example by transporting the NCR peptide signals into the bacterial cytoplasm and so allowing the peptides to interfere with the bacterial cell cycle machinery and other intracellular targets[6,7]. Our phenotypic analyses of the *bclA*/*DD-CPase1* double mutants in the *Bradyrhizobium* strains USDA110 and ORS285 are in favor of the first option; they suggest that the sole role of BclA is the protection against the NCR stress. The abnormal and excessive differentiation provoked by the *DD-CPase1* mutation uncovered the differentiation potential of the *bclA* mutants. This interpretation of our data is in agreement with our previous observation that *in vitro* the *bacA* mutant of *S. meliloti* responds similarly as the wild type strain to NCR treatment by elongation and DNA amplification at peptide concentrations below their toxic level[8]. This implies that the NCR peptides are perceived at the outer or inner membrane, activating a signaling process that induces the differentiation process or alternatively, that NCR peptides can enter in a BclA/BacA independent manner the bacterial cells to reach their cytosolic targets for differentiation.

Methods

Bacterial growth conditions. *Bradyrhizobium* spp. strains were grown in yeast mannitol[45] (YM) or buffered nodulation medium for bacteria (BNM-B)[9] at 30 °C, *E. coli* and *S. meliloti* strains in Luria-Bertani (LB) medium at 37 °C or 30 °C respectively. Antibiotics for strain and plasmid selection were used at the following concentrations: streptomycin (500 µg/ml); spectinomycin (100 µg/ml); chloramphenicol (12.5 µg/ml); gentamycin (15 or 50 µg/ml); carbenicilin (50 µg/ml); tetracycline (10 µg/ml); kanamycin (50 or 200 µg/ml); neomycin (120 µg/ml); and nalidixic acid (25 µg/ml).

Construction of bacterial mutants and complemented strains. Standard molecular biology techniques were used for all cloning work. The construction of the *B. diazoefficiens* USDA110 *bclA* deletion mutant (USDA110*ΔbclA*) was performed with the same strategy as the previously described mutant in *Bradyrhizobium* strain ORS285[9]. The following primers were used for the amplification of the upstream and downstream region of the *bclA* gene: GATAGAAAGCTTAAGCGTCCGGTGGTCACCGTCACCTG (forward, upstream region); GTCTTGCGCCGGATCCAGAGCTCTGTCTCCTGAGGGGATG (reverse, upstream region); ACAGAGCTCTGGATCCGGCGCAAGACGATCGCTATCGTAG (forward, downstream region); GATAGAACTAGTGTTCGGCACCTCGGACGCCTTCTAC (reverse, downstream region).

The mutant strains ORS285*ΔbclA*, ORS285*ΔDD-CPase1* and USDA110*ΔDD-CPase1* were obtained from previous studies[9,29]. The double mutants ORS285*ΔbclADD-CPase1* and USDA110 *ΔbclADD-CPase1* were constructed from the corresponding single mutant in *bclA* and using the respective *DD-CPase1* pVO155-npt2-GFP mutagenesis plasmids which were introduced by triparental mating, as described[29].

The *bclA* gene of *B. diazoefficiens* USDA110 was amplified by PCR using the following primer couple: TTATCGTCTAGACCCTCAGGAGACAGAGCTCTGTGAAG (forward) and CATGATGGATCCGATCGTCTT GCGCCTCAGCGCGCCACGGTC (reverse). The PCR fragment was cloned into the *Xba*I and *Bam*HI sites of plasmid pRF771[46]. The resulting plasmid was introduced into *S. meliloti* Sm1021*ΔbacA* by tri-parental conjugation and into *E. coli* BW25113*ΔsbmA* by electroporation. The same USDA110 *bclA*-encoding PCR fragment was also introduced in plasmid pMG103[47]. First, the *trp* promoter was amplified with primers GTGCCGAATTCGGCAAATATACTGAAATAGGTGTTG (forward) and GAGTGCATGCGGTA CCGGATCCATGGAATCTAGATTTAAAGTACTTCGAA (reverse) and introduced into the *Eco*RI and *Sph*I sites of pMG103. The *bclA* PCR product was subsequently cloned in the *Xba*I and *Bam*HI sites of this pMG103 derivative. The resulting construct was introduced into *Bradyrhizobium* sp. ORS285*ΔbclA* by electroporation.

Peptide sensitivity assays, single-cell NCR peptide uptake assay and flow cytometry. All sensitivity assays were as described before[9]. The peptide uptake assay by flow cytometry was as described[48].

Real-Time Quantitative PCR Expression Analyses. The relative expression level of *NCR* genes was determined by RT-qPCR using methods for RNA extraction, cDNA synthesis, quantitative PCR, normalization and with primers previously described[4].

Plant growth and nodulation. The soybean variety used in this study was *Glycine max* Williams 82. Seeds were cleaned with 100% ethanol for 30 sec and then surface-sterilized with bleach for 15 min. *A. indica* and *A. afraspera* seeds were surface sterilized using sulfuric acid (30 min) and bleach (30 min). Sterilized seeds were then transferred at 28 °C in the darkness on agar plates (0.8%) with tap water for germination. One-day old seedlings were transferred to test tubes containing liquid BNM[49]. Seedlings were grown at 28 °C with a 16 h light regime and 70% humidity. Seven days after germination, each seedling was inoculated with 1 ml of a bacterial suspension adjusted to an $OD_{600} = 0.1$.

Symbiotic analyses. Nitrogen-fixation assays were performed with ten plants per condition at 14 days post-inoculation, as previously described[9]. Plant biomass production was determined from the fresh weight of shoots that were removed from whole plants. The nitrogen and carbon content was determined on pools of three de-nodulated whole plants that were dried in an oven at 60 °C for 48 h. The dried samples were analyzed with an Isoprime element analyzer (Elementar).

For microscopy analysis, nodules were harvested, embedded with agarose (6%), and then freshly sectioned with a Leica VT1200S vibratome (Leica Microsystems GmbH) into 50 µm tissue slices. Slices were incubated for 20 min in Live/Dead BacLight (Molecular Probes) staining solution in Tris buffer (50 mM), pH 7.0, containing 0.01% CalcofluorWhiteM2R (Sigma). Sections were washed to remove excess of dye and observed using a Leica TCS SP8x confocal microscope.

Toluidine blue staining of thin sections was made as follows. Nodules were fixed in 1% glutaraldehyde/4% paraformaldehyde. After washing, nodules were dehydrated in ethanol series and embedded in Technovit 7100 resin (KulzerHistoTechnik) according to the manufacturer's instructions. Five µm sections were obtained with a Leica RM2155 microtome and stained with toluidine blue (0.005%). Bright field microscopy was performed with an Eclipse 80i microscope (Nikon).

Bacteroid purification and flow cytometry analysis. Bacteroids were purified as described[24]. Bacteria were then fixed by heat treatment (70 °C, 10 min) and stained by DAPI (20 µg/ml) before flow cytometry analysis using a MoFlo ASTRIOS flow cytometer (Beckman Coulter). The bacterial DNA content was assessed by the DAPI fluorescence with a 355-nm laser line. Each single event was recorded and analyzed with the Summit 6.2 software (Beckman Coulter).

Data Availability. All data generated or analyzed during this study are included in this published article (and its Supplementary Information files).

References

1. Kondorosi, E., Mergaert, P. & Kereszt, A. A paradigm for endosymbiotic life: cell differentiation of *Rhizobium* bacteria provoked by host plant factors. *Annu. Rev. Microbiol.* **67**, 611–28 (2013).
2. Alunni, B. & Gourion, B. Terminal bacteroid differentiation in the legume–rhizobium symbiosis: nodule-specific cysteine-rich peptides and beyond. *New Phytol.* **211**, 411–417 (2016).
3. Van de Velde, W. *et al.* Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* **327**, 1122–1126 (2010).

4. Czernic, P. *et al.* 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 (2015).
5. Montiel, J. *et al.* Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proc. Natl. Acad. Sci. USA* **114**, 5041–5046 (2017).
6. Farkas, A. *et al. Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proc. Natl. Acad. Sci. USA* **111**, 5183–5188 (2014).
7. Penterman, J. *et al.* Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. *Proc. Natl. Acad. Sci. USA* **111**, 3561–3566 (2014).
8. Haag, A. F. *et al.* Protection of *Sinorhizobium* against host cysteine-rich antimicrobial peptides is critical for symbiosis. *PLoS Biol.* **9**, e1001169, doi:10.1371/journal.pbio.1001169 (2011).
9. Guefrachi, I. *et al. Bradyrhizobium* BclA is a peptide transporter required for bacterial differentiation in symbiosis with *Aeschynomene* legumes. *Mol. Plant. Microbe. Interact.* **28**, 1155–1166 (2015).
10. Runti, G. *et al.* Functional characterization of SbmA, a bacterial inner membrane transporter required for importing the antimicrobial peptide Bac7(1–35). *J. Bacteriol.* **195**, 5343–5351 (2013).
11. Laviña, M., Pugsley, A. P. & Moreno, F. Identification, mapping, cloning and characterization of a gene (*sbmA*) required for microcin B17 action on *Escherichia coli* K12. *J. Gen. Microbiol.* **132**, 1685–1693 (1986).
12. Yorgey, P. *et al.* Posttranslational modifications in microcin B17 define an additional class of DNA gyrase inhibitor. *Proc. Natl. Acad. Sci. USA.* **91**, 4519–4523 (1994).
13. Salomón, R. A. & Farías, R. N. The peptide antibiotic microcin 25 is imported through the TonB pathway and the SbmA protein. *J. Bacteriol.* **177**, 3323–3325 (1995).
14. Mattiuzzo, M. *et al.* Role of the *Escherichia coli* SbmA in the antimicrobial activity of proline-rich peptides. *Mol. Microbiol.* **66**, 151–163 (2007).
15. Pränting, M., Negrea, A., Rhen, M. & Andersson, D. I. Mechanism and fitness costs of PR-39 resistance in *Salmonella enterica* serovar typhimurium LT2. *Antimicrob. Agents Chemother.* **52**, 2734–2741 (2008).
16. Marlow, V. L. *et al.* Essential role for the BacA protein in the uptake of a truncated eukaryotic peptide in *Sinorhizobium meliloti*. *J. Bacteriol.* **191**, 1519–1527 (2009).
17. Wehmeier, S. *et al.* Internalization of a thiazole-modified peptide in *Sinorhizobium meliloti* occurs by BacA-dependent and -independent mechanisms. *Microbiology* **156**, 2702–2713 (2010).
18. Ghosal, A., Vitali, A., Stach, J. E. M. & Nielsen, P. E. Role of SbmA in the uptake of Peptide Nucleic Acid (PNA) -peptide conjugates in *E. coli*. *ACS Chem. Biol.* **8**, 360–367 (2012).
19. Puckett, S. E. *et al.* Bacterial resistance to antisense peptide phosphorodiamidate morpholino oligomers. *Antimicrob. Agents Chemother.* **56**, 6147–6153 (2012).
20. Arnold, M. F. F. *et al.* Partial complementation of *Sinorhizobium meliloti bacA* mutant phenotypes by the *Mycobacterium tuberculosis* BacA protein. *J. Bacteriol.* **195**, 389–398 (2013).
21. Narayanan, S. *et al.* Mechanism of *Escherichia coli* resistance to pyrrhocoricin. *Antimicrob. Agents Chemother.* **58**, 2754–2762 (2014).
22. Guefrachi, I., Verly, C., Kondorosi, E., Alunni, B., & Mergaert, P. Role of the bacterial BacA ABC-transporter in chronic infection of nodule cells by *Rhizobium* bacteria. In *Biological Nitrogen Fixation* (ed. de Bruijn, F. J.) 315-324 (John Wiley & Sons, Inc., 2015).
23. Mergaert, P. *et al.* 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 (2003).
24. Mergaert, P. *et al.* Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. *Proc. Natl. Acad. Sci. USA* **103**, 5230–5235 (2006).
25. Maruya, J. & Saeki, K. 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 Cell Physiol.* **51**, 1443–1452 (2010).
26. Crespo-Rivas, J. C. *et al. 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 (2016).
27. Ardissone, S. *et al.* Role of BacA in lipopolysaccharide synthesis, peptide transport, and nodulation by *Rhizobium* sp. strain NGR234. *J. Bacteriol.* **193**, 2218–2228 (2011).
28. Karunakaran, R. *et al.* BacA is essential for bacteroid development in nodules of galegoid, but not phaseoloid, legumes. *J. Bacteriol.* **192**, 2920–2928 (2010).
29. Gully, D. *et al.* A peptidoglycan-remodeling enzyme is critical for bacteroid differentiation in *Bradyrhizobium* during legume symbiosis. *Mol. Plant. Microbe. Interact.* **29**, 447–457 (2016).
30. Mardirossian, M. *et al.* The host antimicrobial peptide $Bac7_{1-35}$ binds to bacterial ribosomal proteins and inhibits protein synthesis. *Chem. Biol.* **21**, 1639–1647 (2014).
31. diCenzo, G. C., Zamani, M., Ludwig, H. N. & Finan, T. M. Heterologous complementation reveals a specialized activity for BacA in the *Medicago – Sinorhizobium meliloti* symbiosis. *Mol. Plant Microbe Interact.* **30**, 312–324 (2017).
32. Renier, A. *et al.* Photosynthetic *Bradyrhizobium* sp. strain ORS285 synthesizes 2-O- methylfucosylated lipochitooligosaccharides for *nod* gene–dependent interaction with *Aeschynomene* plants. *Mol. Plant Microbe Interact.* **24**, 1440–1447 (2011).
33. Kannenberg, E. L., Perzl, M. & Härtner, T. The occurrence of hopanoid lipids in *Bradyrhizobium* bacteria. *FEMS Microbiol. Lett.* **127**, 255–261 (1995).
34. Schmerk, C. L., Bernards, M. A. & Valvano, M. A. Hopanoid production is required for low-pH tolerance, antimicrobial resistance, and motility in *Burkholderia cenocepacia*. *J. Bacteriol.* **193**, 6712–6723 (2011).
35. Malott, R. J., Steen-Kinnaird, B. R., Lee, T. D. & Speert, D. P. Identification of hopanoid biosynthesis genes involved in polymyxin resistance in *Burkholderia multivorans*. *Antimicrob. Agents Chemother.* **56**, 464–471 (2012).
36. Kulkarni, G. *et al.* Specific hopanoid classes differentially affect free-living and symbiotic states of *Bradyrhizobium diazoefficiens*. *MBio* **6**, e01251–15, doi:10.1128/mBio.01251-15 (2015).
37. Giraud, E. *et al.* Legumes symbioses: absence of *nod* genes in photosynthetic bradyrhizobia. *Science* **316**, 1307–1312 (2007).
38. Mornico, D. *et al.* Comparative genomics of *Aeschynomene* symbionts: insights into the ecological lifestyle of nod-independent photosynthetic bradyrhizobia. *Genes* **3**, 35–61, doi:10.3390/genes3010035 (2012).
39. Kazmierczak, T. *et al.* Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium* strains. *Mol. Plant Microbe Interact.* **30**, 399–409 (2017).
40. Oono, R. & Denison, R. F. Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol.* **154**, 1541–1548 (2010).
41. Kereszt, A., Mergaert, P. & Kondorosi, E. Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims? *Mol. Plant-Microbe Interact.* **24**, 1300–1309 (2011).
42. Sen, D., Weaver, R. W. & Bal, A. K. Structure and organization of effective peanut and cowpea root nodules induced by rhizobial strain 32H1. *J. Exp. Bot.* **37**, 356–363 (1986).
43. Yang, S. *et al.* Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proc. Natl. Acad. Sci. USA* **114**, 6848–6853 (2017).
44. Wang, Q. *et al.* Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. *Proc. Natl. Acad. Sci. USA* **114**, 6854–6859 (2017).

45. Giraud, E., Hannibal, L., Fardoux, J., Verméglio, A. & Dreyfus, B. Effect of *Bradyrhizobium* photosynthesis on stem nodulation of *Aeschynomene sensitiva*. *Proc. Natl. Acad. Sci. USA* **97**, 14795–14800 (2000).
46. Wells, D. H. & Long, S. R. The *Sinorhizobium meliloti* stringent response affects multiple aspects of symbiosis. *Mol Microbiol.* **43**, 1115–1127 (2002).
47. Inui, M., Roh, J. H., Zahn, K. & Yukawa, H. Sequence analysis of the cryptic plasmid pMG101 from *Rhodopseudomonas palustris* and construction of stable cloning vectors. *Appl Environ Microbiol.* **66**, 54–63 (2000).
48. Benincasa, M. *et al.* Single cell flow cytometry assay for peptide uptake by bacteria. *Bio-Protocol* **6**, e2038, doi:10.21769/BioProtoc.2038 (2016).
49. Ehrhardt, D. W. *et al.* Depolarization of alfalfa root hair membrane potential by *Rhizobium meliloti* Nod factors. **256**, 998–1000 (1992).

Acknowledgements

We thank Christien Merrifield (Institute for Integrative Biology of the Cell) for comments on the manuscript. Q.B. and F.L. were supported by PhD fellowships from the Université Paris-Sud and I.G. by PhD fellowships from the Tunisian Ministry of Higher Education and Scientific Research, the International Relations Office of Université Paris-Sud and the Agence Universitaire de la Francophonie (AUF). 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 Science' (ANR-11-IDEX-0003-02). This work was funded by the Agence Nationale de la Recherche, grant n° ANR-13-BSV7-0013.

Author Contributions

Q.B., I.G., E.G. and P.M. conceived the experiments; Q.B., I.G., D.G., F.L., O.P., J.F. and C.C. performed the experiments; Q.B., I.G., T.T., B.A., E.G. and P.M. analyzed the data. Q.B. and P.M. wrote the paper and all authors reviewed and corrected the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-08830-0

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017

Integrated roles of BclA and DD-carboxypeptidase 1 in *Bradyrhizobium* differentiation within NCR-producing and NCR-lacking root nodules

Quentin Barrière, Ibtissem Guefrachi, Djamel Gully, Florian Lamouche, Olivier Pierre, Joël Fardoux, Clémence Chaintreuil, Benoît Alunni, Tatiana Timchenko, Eric Giraud, Peter Mergaert

Supplementary information

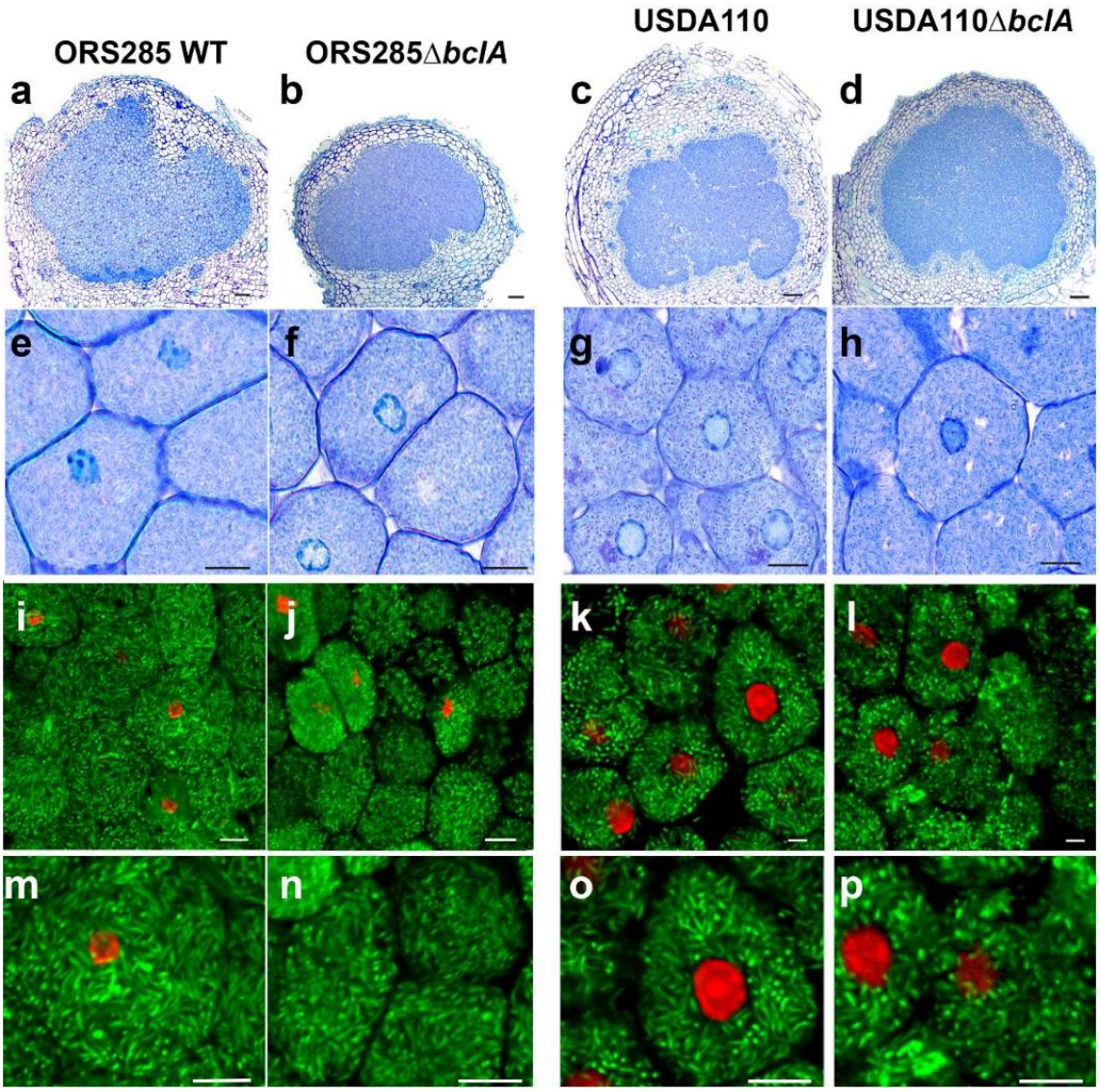


Fig. S1. Histology of *Aeschynomene afraspera* nodules infected with strains ORS285 or USDA110 and their *bclA* mutants. (**a-h**) Light microscopy of toluidine blue stained thin sections. (**i-p**) Confocal microscopy of fresh nodule sections stained with Live/Dead BacLight. Scale bars are 100 µm (a-d) or 10 µm (e-p).

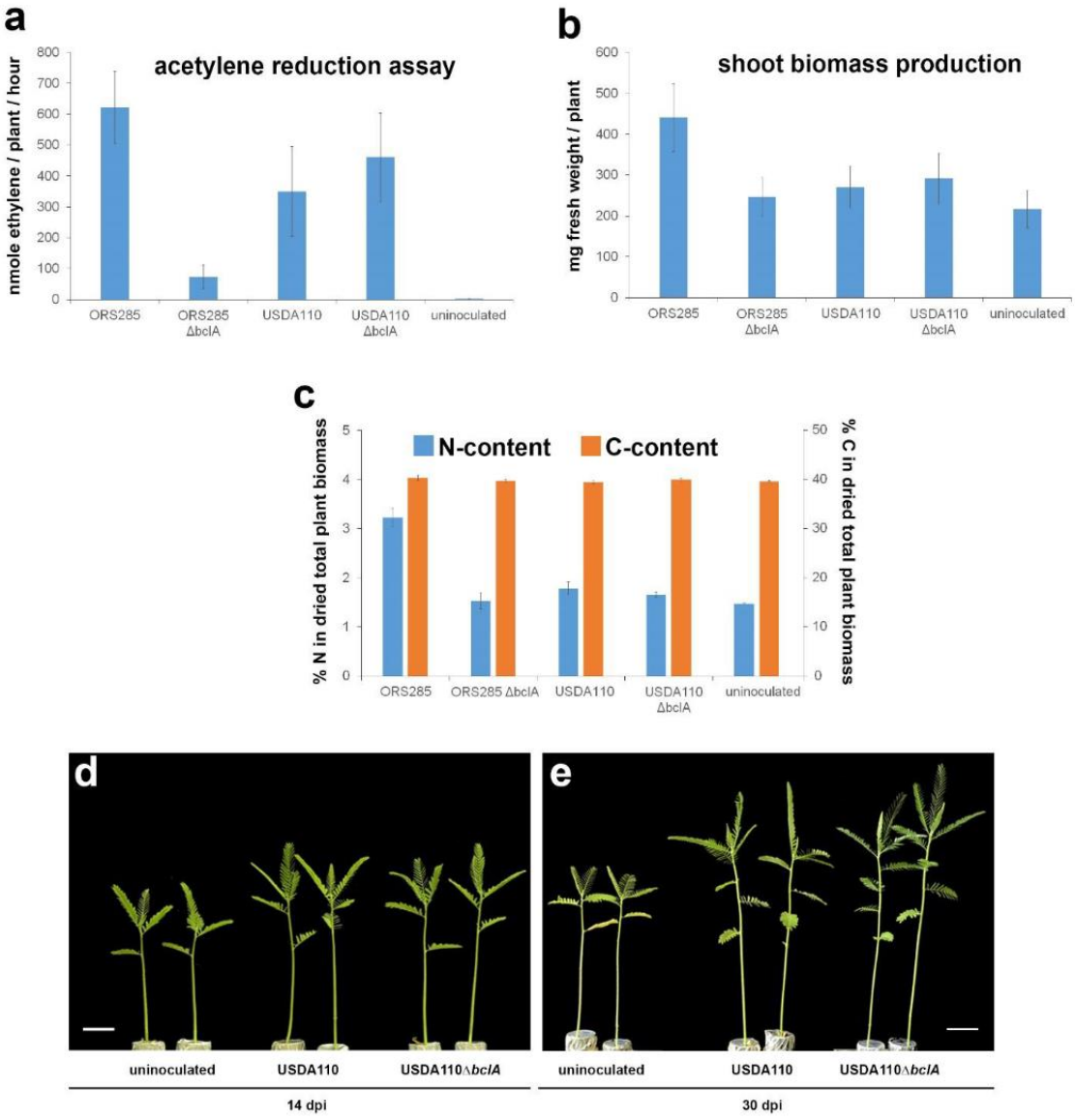


Fig. S2. Nitrogen fixation, plant growth and nitrogen content in *Aeschynomene afraspera* infected with strains ORS285, USDA110 and their *bclA* mutants. (**a**) Acetylene reduction assay of whole plants at 21 dpi. n=10 and error bars are standard deviations. (**b**) Shoot fresh weight of plants at 21 dpi. n=10 and error bars are standard deviations. (**c**) Nitrogen and carbon content of plants expressed as % of dry mass of de-nodulated plants. Each analysis was made on pools of three plants. n=2 or 3 and error bars are standard deviations. (**d**,**e**) Plant growth of uninoculated *A. afraspera* plants or plants inoculated with *B. diazoefficiens* USDA110 wild type or its *bclA* mutant at 14 dpi (d) or 30 dpi (e). Scale bars are 2 cm.

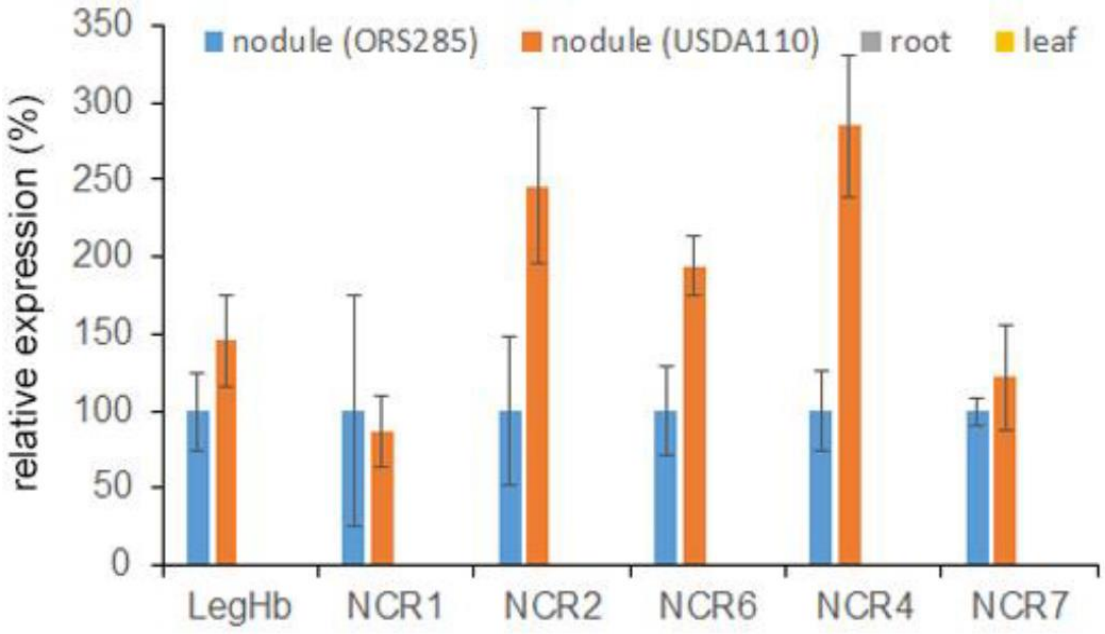


Fig. S3. Relative expression of 5 *NCR* genes and leghemoglobin in *A. afraspera* roots, leaves and nodules infected with *B. diazoefficiens* USDA110. The expression of each gene was normalized by the constitutive elongation factor 1α gene and set at 100% in the *A. afraspera* nodules infected with *Bradyrhizobium* strain ORS285. The expression of all six genes was below detection limit in the root and leave samples. Error bars are standard deviations of three biological repeats.

3. The biotroph *Agrobacterium tumefaciens* thrives in tumors by exploiting a wide spectrum of plant host metabolites

Almudena Gonzalez-Mula, Joy Lachat, Léo Mathias, Delphine Naquin, Florian Lamouche, Peter Mergaert et Denis Faure

Publié dans la revue *New Phytologist* (2019). DOI: https://doi.org/10.1111/nph.15598

New Phytologist

Research

The biotroph *Agrobacterium tumefaciens* thrives in tumors by exploiting a wide spectrum of plant host metabolites

Almudena Gonzalez-Mula, Joy Lachat, Léo Mathias, Delphine Naquin, Florian Lamouche, Peter Mergaert and Denis Faure

Institute for Integrative Biology of the Cell (I2BC), CNRS CEA University Paris-Sud, University Paris-Saclay, Gif-sur-Yvette F-91190, France

Author for correspondence:
Denis Faure
Tel: +33 169823498
Email: *denis.faure@i2bc.paris-saclay.fr*

Received: *22 August 2018*
Accepted: *3 November 2018*

New Phytologist (2019) **222:** 455–467
doi: 10.1111/nph.15598

Key words: biotroph, ecological niche, fitness, plant pathogen, transcriptomics, transposon-sequencing (Tn-seq).

Summary

- *Agrobacterium tumefaciens* is a niche-constructing biotroph that exploits host plant metabolites.
- We combined metabolomics, transposon-sequencing (Tn-seq), transcriptomics, and reverse genetics to characterize *A. tumefaciens* pathways involved in the exploitation of resources from the *Solanum lycopersicum* host plant.
- Metabolomics of healthy stems and plant tumors revealed the common (e.g. sucrose, glutamate) and enriched (e.g. opines, γ-aminobutyric acid (GABA), γ-hydroxybutyric acid (GHB), pyruvate) metabolites that *A. tumefaciens* could use as nutrients. Tn-seq and transcriptomics pinpointed the genes that are crucial and/or upregulated when the pathogen grew on either sucrose (*pgi*, *kdgA*, *pycA*, *cisY*) or GHB (*blcAB*, *pckA*, *eno*, *gpsA*) as a carbon source. While sucrose assimilation involved the Entner–Doudoroff and tricarboxylic acid (TCA) pathways, GHB degradation required the *blc* genes, TCA cycle, and gluconeogenesis. The tumor-enriched metabolite pyruvate is at the node connecting these pathways. Using reverse genetics, we showed that the *blc*, *pckA*, and *pycA* loci were important for aggressiveness (tumor weight), proliferation (bacterial charge), and/or fitness (competition between the constructed mutants and wild-type) of *A. tumefaciens* in plant tumors.
- This work highlighted how a biotroph mobilizes its central metabolism for exploiting a wide diversity of resources in a plant host. It further shows the complementarity of functional genome-wide scans by transcriptomics and Tn-seq to decipher the lifestyle of a plant pathogen.

Introduction

Hosts and microbes evolved a wide spectrum of biological interactions, ranging from pathogenesis to symbiosis. To succeed in their lifestyle, host-interacting microbes are able to escape host defense, overcome competition with other microbiota members, and exploit nutrients available in the hosts. Ecological niche construction, which ensures a preferred access to host-derived resources, represents a recurrent strategy in pathogens and symbionts (Kylafis & Loreau, 2011; McNally & Brown, 2015; Martin *et al.*, 2017; Poole *et al.*, 2018). Identifying metabolic pathways involved in the exploitation of resources and evaluating their involvement in the fitness of microbes represent important issues in ecology and evolution for understanding adaptation of microbes to the hosts, with applied perspectives in plant, animal, and human health.

Different strategies have emerged to identify the microbial pathways involved in resource exploitation. They basically employ a two-step methodology. The first step is the identification of candidate genes and pathways by different genome-wide scans (functional screening of individual mutants, transcriptomics, genomics, genome wide association, etc.) in microbes that exploit a given resource using, in some instances, a comparison with microbes that do not exploit it. The second step is the validation of a fitness trait by confronting microbes carrying allelic variation (natural or constructed variants) in those candidate genes and pathways. Because of its relative simplicity, transposon sequencing (Tn-seq), which combines transposon insertional mutagenesis with massively parallel sequencing of the transposon insertion sites in transposon mutant populations grown in control and test conditions, seemed an attractive approach to examine ecologically important genes and pathways in prokaryotic and eukaryotic microbes (van Opijnen & Camilli, 2013). In this study, we combined plant metabolomics and two functional genome-wide scans (transcriptomics and Tn-seq) for identifying genes and pathways involved in the exploitation of the *Solanum lycopersicum* host by the biotrophic pathogen *Agrobacterium tumefaciens*.

A. tumefaciens is a niche-constructing pathogen that genetically modifies the plant host genome by transferring a part (the transfer DNA (T-DNA)) of its virulence Ti plasmid (Barton *et al.*, 2018; Dessaux & Faure, 2018). When expressed into the plant cell nucleus, the T-DNA genes divert the host hormonal and

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

metabolic pathways to provoke the development of galls or plant tumors (Deeken *et al.*, 2006). In previous work, we paid attention to specific metabolites, the opines, that accumulate in the *A. tumefaciens*-infected plant tumors (Lang *et al.*, 2014; El Sahili *et al.*, 2015; Marty *et al.*, 2016; Tannières *et al.*, 2017; Lang *et al.*, 2017; Vigouroux *et al.*, 2017). Opines, such as agrocinopines, mannopine, nopaline, and octopine, result from the condensation of sugars and amino and organic acids (Dessaux *et al.*, 1993). According to chemical and genome databases, the opines synthesized by *Agrobacterium* T-DNA-encoded enzymes are almost exclusively produced by host plants infected by *A. tumefaciens*. To our knowledge, the only reported exception is the opine octopine that is also produced in the muscle of the marine animal octopus (Fields *et al.*, 1976). The biosynthesis of opines in the chimeric plant cells expressing bacterial T-DNA could be considered as a biological innovation resulting from the holobiont assembly (Faure *et al.*, 2018). In the cases of nopaline and octopine, we showed that opine assimilation confers a selective advantage when *A. tumefaciens* populations colonize the plant tumors (Lang *et al.*, 2014, 2017; Vigouroux *et al.*, 2017). Aside from opines, diverse metabolites accumulate in the tumors on *Arabidopsis thaliana* and *Brassica rapa* (Deeken *et al.*, 2006; Simoh *et al.*, 2009; Lang *et al.*, 2016), but their contribution to *A. tumefaciens* fitness and proliferation is poorly documented. Recently, the transcriptome of *A. tumefaciens* C58 living in *A. thaliana* tumors highlighted considerable changes in gene expression profile compared with a culture in a synthetic medium (González-Mula *et al.*, 2018). In addition to the opines, the transcriptomic data suggested the exploitation of a wide diversity of resources by *A. tumefaciens*, but direct evidence of the contribution of these different metabolites to the *Agrobacterium* fitness in plant tumors was still missing.

In this work, metabolomics indeed revealed the presence of a wide spectrum of potential resources in *S. lycopersicum* tumors, including metabolites that were enriched compared with uninfected stems. We combined Tn-seq and transcriptomics to investigate the *A. tumefaciens* pathways for exploiting the three metabolites sucrose, γ-hydroxybutyrate, and γ-aminobutyrate that accumulated at different levels in plant tumors. Finally, we used reverse genetics and host plant infections to measure the aggressiveness, proliferation, and competitive fitness conferred by assimilation of these metabolites when *A. tumefaciens* colonized the plant tumor niche. We showed that the ecological success of the *A. tumefaciens* biotroph resulted from its capacity to exploit a wider spectrum of host metabolites than the sole opines. This work also highlighted the strength and limits of Tn-seq and transcriptomics to decipher the microbial genetic determinants that are involved in ecological niche exploitation.

Materials and Methods

Bacterial strains and culture conditions

We used *A. tumefaciens* C58, the genome of which was sequenced in 2001 (Goodner *et al.*, 2001; Wood *et al.*, 2001). The kanamycin (Km)-resistance and gentamicin (Gm)-resistance cassettes (Dennis & Zylstra, 1998) were used for the construction of the knockout (KO) mutants. The *atu0035* (*pckA*), *atu2726* (*pycA*), *atu3706*, and *atu4761* genes were cloned into the pGEM-T Easy vector (Promega), and the mutated alleles were created by inserting an antibiotic-resistance cassette in a unique restriction site of the open reading frame. The constructed plasmids were electroporated in *A. tumefaciens* C58. Marker exchange was selected using Gm or Km resistance and verified by PCR. Previously constructed *A. tumefaciens* C58 mutants were also used in this study: the derivatives C107-Gm and C107-Km in which the Gm and Km cassettes were cloned in a noncoding region of the Ti plasmid (Haudecoeur *et al.*, 2009a) and the Δ*blcRABC* mutant in which the *blcRABC* operon was replaced by the Km-resistance cassette (Carlier *et al.*, 2004).

A. tumefaciens was cultivated at 28°C in TY medium (Bacto tryptone, 5 g l^{-1}; yeast extract, 3 g l^{-1}; agar, 15 g l^{-1}) or *Agrobacterium* broth (AB) minimal medium (dipotassium hydrogen phosphate, 3 g l^{-1}; sodium dihydrogen phosphate, 1 g l^{-1}; magnesium sulfate heptahydrate, 0.3 g l^{-1}; potassium chloride, 0.15 g l^{-1}; calcium chloride, 0.01 g l^{-1}; ferrous sulfate heptahydrate, 2.5 mg l^{-1}; pH 7) (Chilton *et al.*, 1974) supplemented with sucrose or γ-hydroxybutyric acid (GHB) at 10 mM as carbon (C) source, and ammonium chloride (NH_4Cl) or γ-aminobutyric acid (GABA) at 20 mM as nitrogen (N) source. *Escherichia coli* MFD*pir* harboring the pSAM_DGm plasmid (Skurnik *et al.*, 2013), auxotroph for diaminopimelic acid, was used as transposon donor for mutagenesis. *E. coli* DH5α was the routine host for cloning. *E. coli* strains were cultivated at 37°C in lysogenic broth modified medium (LBm; 10 g l^{-1} peptone, 5 g l^{-1} yeast extract, sodium chloride (NaCl) 5 g l^{-1}). Media were supplemented when appropriate with Gm (25 μg ml^{-1}), ampicillin (50 μg ml^{-1}), rifampicin (100 μg ml^{-1}), and diaminopimelic acid (300 μg ml^{-1}).

Plant culture, metabolomics, and infection assays

S. lycopersicum (Dona hybrid F_1, Vilmorin, France) plants were cultivated in a glasshouse under long day conditions and controlled temperature (24–26°C). Four-week-old plants were incised with a scalpel between the first and second nodes and infected by *c.* 10^7 *A. tumefaciens* cells as described by Planamente *et al.* (2010). Plant tumors were collected 4 wk after infection.

For plant metabolomics, tumors and wounded but not infected stems were directly frozen in liquid N_2, crushed, extracted and analyzed by gas chromatography–time of flight mass spectrometry (GC–TOF-MS) at the *Plateforme de Chimie du Végétal* (Versailles, France). The method was previously described in detail by Lang *et al.* (2016). Approx. 150 compounds were searched and 130 compounds, including the opines nopaline and agrocinopine A, and GABA and GHB were detected and quantified in three biological replicates of plant tumors and uninfected stems.

For virulence and fitness assays, eight plant tumors were crushed in a 0.8% NaCl solution to recover the agrobacteria, which were then spotted onto selective agar media to enumerate colony forming units (CFU). In the case of mixed infections, the

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

New Phytologist

Research 457

proportions of the genotypes (wild-type and KO alleles) were measured by testing *c.* 96 CFU. Using appropriate primers (Supporting Information Table S1), length of the PCR products distinguished wild-type allele from the KO-alleles in which the resistance gene cassette was inserted. This permitted calculation of competitive index values as previously described (Macho *et al.*, 2010). Two independent assays (eight plants for each of the assays) were carried out for each virulence and fitness assays. A Mann–Whitney test was used to analyze the values from the two independent experiments (the null hypothesis postulates that both experiments were comparable). If no difference was detected, the values were pooled and a nonparametric Kruskal–Wallis test ($P < 0.05$) coupled with a post-hoc Tukey test ($P < 0.05$) was performed.

Transposon library construction and use

A. tumefaciens C58 was mutagenized using a *Himar1* mariner transposon carrying a Gm resistance cassette. The pSAM_DGm plasmid donor *E. coli* MFDpir and *A. tumefaciens* C58 rifampicin-resistant recipient were cultivated separately: *A. tumefaciens* C58 Rif^R overnight in TY medium and *E. coli* MFDpir (pSAM_DGm) for 4 h in LBm supplemented with 300 μg ml^{-1} diaminopimelic acid. Both cultures were centrifuged and adjusted to 1 unit of OD_{600}. Equivalent volumes (0.4 ml) of cell suspensions were mixed, centrifuged, and suspended in TY with diaminopimelic acid. The cell mixture (0.2 ml) was deposited on a nitrocellulose filter (0.45 μm diameter, Millipore) on a TY agar plate and incubated overnight at 28°C. Bacterial cells were removed from the filter, suspended in 0.8% NaCl solution, and then plated on TY medium supplemented with rifampicin and Gm. Serial dilutions and plating were performed to determine the number of mutants obtained. After 72 h of incubation, mutants were collected. The mutant population was homogenized, aliquoted, and stored at −80°C in 25% (v/v) glycerol.

Four aliquots of the *A. tumefaciens* mutant library were thawed and cultured in liquid TY medium (4 h at 28°C) to revive them. Bacteria were washed twice with 0.8% NaCl solution and used to inoculate AB medium (10 ml) at an initial OD_{600} of 0.05. AB medium was supplemented with three combinations of C and N sources: sucrose–ammonium (NH_4), GHB–NH_4, and sucrose–GABA. After growth at 28°C for 24 h, bacterial cells were centrifuged and stored at −20°C for further DNA manipulation.

Transposon library sequencing and ARTIST analysis

Genomic DNA of mutant populations grown in TY medium or AB medium with the C and N sources tested was extracted using the DNeasy Blood & Tissue Kit (Qiagen). DNA (2 μg) was digested with the *Mme*I type II restriction–modification enzyme (BioLabs, Evry, France) for 1 h at 37°C. Digested DNA was incubated for 1 h at 37°C with FastAP thermosensitive alkaline phosphatase (Thermo Scientific, Waltham, MA, USA) followed by enzyme inactivation by heating at 75°C for 5 min. Digested DNA samples were purified using a QIAquick PCR purification kit (Qiagen) and were ligated to the p-adapters (Table S1) in the presence of Thermo Scientific T4 DNA ligase (16 h at 16°C). The p-adaptors contain a five-nucleotide long barcode that is specific for each experiment. The ligation products were used as templates to perform a PCR amplification with Illumina-primers P7 and P5 (Table S1). The PCR products of *c.* 130 base pairs, which contain the transposon insertion site, were separated on agarose gel and purified with the QIAquick gel extraction kit (Qiagen). The final samples were mixed in equimolar amounts and sequenced on an Illumina NextSeq 500 instrument (Illumina, San Diego, CA, USA), in a paired-end 2 × 75 run at the I2BC-sequencing platform (Gif-sur-Yvette, France).

The experiment-specific barcodes enabled the attribution of each sequence read to the corresponding experiment. The data generated were demultiplexed using BCL2FASTQ2 v.2.15.0 (Illumina) and FASTX-TOOLKIT software (http://hannonlab.cshl.edu/fastx_toolkit/). Only read 1 from each sequenced fragment has been used. The 3′ transposon sequence was trimmed using TRIMMOMATIC (Bolger *et al.*, 2014), and reads with a length of 75 nucleotides were removed (reads without the transposon insertion). After the trimming step, reads with a length between 19 and 23 bp were reverse-complemented and only the reads starting with TA were mapped using BOWTIE (BOWTIE-1.1.2) (Langmead *et al.*, 2009) to the genome of *A. tumefaciens* C58. The *.bam output files were sorted with SAMTOOLS (http://www.htslib.org/). FEATURECOUNTS (Liao *et al.*, 2014) was used to evaluate the number of reads by gene or by coding sequence.

The mapping results (*.bam files) were analyzed by the ARTIST pipeline (Pritchard *et al.*, 2014) using MATLAB software (The MathWorks, Natick, MA, USA). ARTIST compares the observed (reads) and predicted numbers of transposons at each of the 115 525 insertion sites (TA dinucleotides) along the *A. tumefaciens* C58 genome. Two different analyses were carried out: EL-ARTIST (essential loci analysis) and CON-ARTIST (conditionally essential loci analysis). EL-ARTIST searches for a nonrandom distribution of transposon insertions in the constructed mutant library in TY medium. Hence, it identifies all loci that are required for an optimal growth in the initial culture condition. A gene is annotated as 'essential' when there are a low number of transposon insertions (reads) or no associated transposon insertions within the entire gene. In EL-ARTIST, 0.03 is the *P*-value threshold for calling a region significantly underrepresented in reads. Then, CON-ARTIST was applied to compare the distribution of transposon insertions between the initial TY culture condition and each of the three AB medium conditions. In CON-ARTIST, 0.01 is the *P*-value cutoff in the Mann–Whitney *U* test for defining genes with significantly different read numbers. This allowed the identification of *A. tumefaciens* genes and pathways that were required for efficient proliferation in the presence of sucrose–NH_4, GHB–NH_4, and sucrose–GABA as nutrients.

Transcriptomics and DESEQ2 analysis

An overnight culture of *A. tumefaciens* C58 was grown in AB medium with sucrose and NH_4 as sources of C and N

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

respectively. This culture was washed twice with NaCl 0.8% and served to inoculate AB medium supplemented with the three different combinations of C and N sources as earlier for the Tn-seq experiments: sucrose–NH_4, GHB–NH_4, and sucrose–GABA. Inoculations (at $OD_{600} = 0.05$) were performed in triplicate. At exponential phase (at $OD_{600} = 0.30$), bacterial cultures were centrifuged and RNA extracted with the MasterPure™ Complete DNA and RNA Purification Kit according to the supplier's instructions. RNA-sequencing (RNA-seq) libraries were constructed using the Ribo-Zero and ScriptSeq-V2 kits (Illumina). Libraries were sequenced on a NextSeq 500 instrument (Illumina) at the I2BC platform (Gif-sur-Yvette, France) using the 75-cycles NextSeq 500 High Output Kit. Count tables have been filtered to retain only genes with a gene count over 1 count per million in half of the samples of the dataset. Normalization and differential analyses were performed using generalized linear models as described in the DESeq2 package (v.1.12.4) (Love *et al.*, 2014). The cutoff chosen for differentially expressed genes are a false discovery rate < 0.01 and a $\log_2$ fold change > 2. RNA-seq data from this article were deposited at https://www.ncbi.nlm.nih.gov/geo/info/linking.html using accession GSE121889, according to Minimum Information About a Microarray Experiment standards.

Gene expression was also measured by quantitative PCR (qPCR) using dedicated primers (Table S1). The cDNA was prepared from 1 μg of bacterial RNA using the RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas, Saint-Remy-les-Chevreuses, France) following the manufacturer's instructions. The qPCRs were performed with a Lightcycler 96 (Roche) apparatus. The data were processed using the $2^{-\Delta\Delta C_T}$ method (Livak & Schmittgen, 2001) and compared with the expression profile acquired from the RNA-seq transcriptome. The internal controls used were the *atu1789* (for the GHB condition) and *nocT* (*atu6027*, for the GABA condition) genes.

Data accessibility

All the data are available in Supporting Information for the article.

Results

Metabolic resources in the plant tumor niche

The abundance of 130 compounds was quantified in uninfected stems of *S. lycopersicum* and tumors induced by *A. tumefaciens* C58 (Fig. 1; Table S2). In plant tumors, the 13 most abundant compounds accounted for 97% of the relative abundance of all the compounds quantified. In decreasing order, these are propanediol, glucose, malate, dehydroascorbate, fructose, phosphate, sucrose, glutamine, glutamate, *myo*-inositol, citrate, asparagine, and aspartate. They were also present in healthy stems at a similar level. Using a threshold fold change value (≥4), 24 compounds were enriched in plant tumors compared with uninfected stems (Fig. 1). As expected, the tumor-enriched compounds encompassed the two opines nopaline and agrocinopine, but some other metabolites were also remarkable. Six metabolites, agrocinopine, *trans*-ferulate, nicotianamine, pyruvate, spermidine, and succinic semialdehyde (SSA), exhibited an enrichment by four to six orders in plant tumors. Their accumulation reflected some well-known characteristics of the plant tumors: obviously, the opine synthesis driven by T-DNA; but also the activation of plant defense as revealed by accumulation of *trans*-ferulate phenolics as antimicrobial compounds and nicotianamine as iron chelator (Deeken *et al.*, 2006; Aznar *et al.*, 2015); a response to abiotic (hypoxia and drought) and biotic stresses, as shown by the accumulation of SSA and spermidine that are related to the GABA pathway (Lang *et al.*, 2016; Podlešáková *et al.*, 2019); and a shift to an anaerobic and heterotrophic metabolism, as suggested by pyruvate accumulation (Deeken *et al.*, 2006).

GABA, SSA, and GHB were all enriched metabolites in plant tumors (Fig. 1). They are metabolically connected and, together with proline, are involved in regulation of the quorum-sensing signal-degrading lactonase BlcC in *A. tumefaciens* (Carlier *et al.*, 2004; Chevrot *et al.*, 2006; Chai *et al.*, 2007; Haudecoeur *et al.*, 2009b; Lang *et al.*, 2016). Noticeably, GABA and proline were the two enriched metabolites that accumulated at the highest concentrations (Fig. 1). As osmoprotectants, proline and sucrose (the latter was not enriched but very abundant) were proposed to contribute to desiccation resistance in plant tumors (Wächter *et al.*, 2003).

In the next part of this work, we focused on compounds other than the opines, and investigated how *A. tumefaciens* could be able to use them as a resource. Because of its role in tumor development, we chose sucrose as a representative of the abundant class of metabolites. We also studied two structurally and functionally related metabolites, GABA and GHB, which were enriched in plant tumors, either at a high concentration (GABA) or at a lower concentration (GHB).

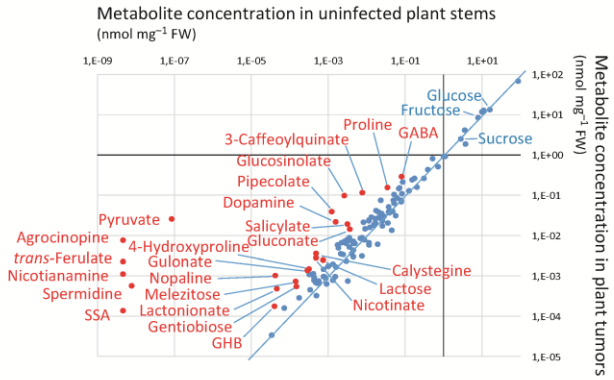


Fig. 1 Metabolome of *Solanum lycopersicum* healthy stems vs *Agrobacterium tumefaciens* C58 tumors. The graphic shows the relative abundance of compounds found in *A. tumefaciens* tumors when compared with uninfected tissue. Red dots indicate metabolites enriched in plant tumors compared with uninfected stems (fold change value ≥ 4); blue dots are the other plant metabolites. GABA, γ-aminobutyric acid; GHB, γ-hydroxybutyric acid; SSA, succinic semialdehyde.

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

New Phytologist

Research 459

Genome-wide mutant library of *A. tumefaciens*

The *A. tumefaciens* C58 genome contained 115 525 TA dinucleotides that are potential insertion sites of the *Himar1* mariner transposon. They are positioned along the circular chromosome (55 348 TA within 2 841 580 bp), linear chromosome (41 503 TA within 2 075 577 bp), and the two plasmids, the pAt (13 084 TA within 542 868 bp) and pTi (5 590 TA within 214 233 bp). From 55 matings between the *E. coli* transposon donor and *A. tumefaciens* C58 recipient, we collected 1.1×10^6 mutant colonies on TY medium supplemented with rifampicin and Gm, hence *c.* 10-fold more than the TA site number in *A. tumefaciens*. After homogenization of all mutant colonies, around 3×10^{10} individuals were kept in each frozen aliquot.

For analyzing the constructed transposon mutant population, four mutant library aliquots were cultivated for 4 h in TY medium. Total DNA was extracted and transposon insertion sites were sequenced. Between 4 million and 8 million filtered reads were obtained for each replicate. When the replicates were compared, transposon distribution revealed a high homogeneity ($r^2 > 0.98$); hence, all the sequencing reads were analyzed together by El-Artist. Mutants in most genes (4730) (Fig. 2a; Table S3.1) were unaffected in their fitness ('nonessential gene' according to the El-Artist classification). Mutants in 513 genes ('essential genes' according to the El-Artist classification) were impaired in their fitness for growth in the rich TY medium (Fig. 2c; Table S3.1). Some other genes (105) (Fig. 2b; Table S3.1) contained a domain in which transposon insertion provoked a decrease of the fitness (genes with an 'essential domain' according to the El-Artist classification). The 513 fitness genes in TY medium represented *c.* 10% of the total genes of *A. tumefaciens* C58, a similar percentage as reported in other bacteria (Christen *et al.*, 2011; DeJesus & Ioerger, 2013). Most of them (428, hence 83%) were located on the circular chromosome (Table S3.1). Using a Tn*5* mutant library, Curtis & Brun (2014) reported 372 essential genes in *A. tumefaciens* C58. Even if the two approaches were different in the choice of transposons, growth condition, library sequencing, and data analysis, most of the Tn*5*-picked essential genes (307 of 372) were also present in the list of the *Himar1*-identified mutants (Table S3; Fig. S1), hence consolidating the two approaches.

The 513 fitness genes were classified according to clusters of orthologous genes (COG) (Tatusov *et al.*, 2000). The functional category most represented was that of translation, ribosomal structure, and biogenesis (Fig. S2; Table S3.1). Genes coding for some ribosomal proteins (*atu1928–atu1951*) exemplified this COG category. Genes involved in the COG category energy production and conversion were also found to be abundant. This is the case of the *nuoABCDEFGHIJKLMN* (*atu1268–atu1283*) operon involved in respiration. Noticeably, the Tn-seq approach revealed some genes that are not essential for cell viability but essential for the maintenance of the *A. tumefaciens* C58 plasmids, such as the operons *repABC* in the At (*atu5000–atu5002*) and pTi (*atu6043–atu6045*) plasmids. This is explained by a transposon insertion in these replicative functions causing the loss of the respective plasmid and, hence, after growth, the disappearance of these mutants in the mutagenized population.

A. tumefaciens key-genes for exploiting sucrose, GHB, and GABA

The transposon mutant population was cultivated in a minimal medium for searching the genes associated with exploitation of either sucrose or GHB as a C source (with NH_4 as a N source) and GABA as a N source (with sucrose as a C source). For each of the bacterial culture replicates, between 3 million and 6 million filtered reads were obtained and analyzed by Artist. Among replicates of a same condition, transposon distribution was highly correlated ($r^2 > 0.93$); hence, reads of a same condition were pooled. By comparing the transposon distribution between the initial growth condition in TY rich medium and the three culture conditions in minimal media, Con-Artist revealed 69, 37, and 47 genes of which mutants were impaired for growth in the presence of sucrose–NH_4, GHB–NH_4, and sucrose–GABA, respectively (Fig. 3; Table S3.2–S3.4). Most of them are involved in amino acid and nucleobase biosynthesis and were shared between the conditions tested. This was expected because cultures in AB minimal medium were compared with an initial culture in TY rich medium. We focused on genes that are specific to each of the growth conditions: there were 28 genes identified in the sucrose–NH_4 condition, 11 in the GHB–NH_4 condition, and nine in the sucrose–GABA condition.

In the presence of sucrose and NH_4 as nutrients, the noticeable genes were *pgi* (*atu0404*) coding for glucose-6-phosphate isomerase, *pycA* (*atu2726*) allowing conversion of pyruvate into oxaloacetate, *cisY* (*atu1392*) for conversion of oxaloacetate into citrate, and the *sdhCDA* (*atu2643–atu2645*) genes for malate conversion and energy production (Table S3.2). These genes are pivotal for the entry of C compounds into the Entner–Doudoroff pathway and the tricarboxylic acid (TCA) cycle. When *A. tumefaciens* grew on GHB–NH_4, key fitness genes were *blcAB* (*atu5137–atu5138*) coding for the conversion of GHB into succinate, as well as *sdhDC*, *pckA* (*atu0035*) and *eno* (*atu1426*) for connecting the TCA cycle and gluconeogenesis, and *gpsA* (*atu2650*) that links gluconeogenesis and lipid biosynthesis (Table S3.3). In the presence of GABA as an N source, we did not identify the expected GABA transaminase key gene that could convert GABA into succinic semialdehyde, probably because of functional redundancy.

A. tumefaciens transcriptomes during growth on sucrose, GHB, and GABA

Using the same minimal media supplemented with sucrose and NH_4, GHB and NH_4, and sucrose and GABA, we produced RNA-seq transcriptomes of *A. tumefaciens* C58 under exponential growth culture condition. RNA-seq trancriptomic data were validated by qPCR assays on a set of nine genes (Fig. S3). In the GHB–NH_4 vs sucrose–NH_4 transcriptome comparison (Fig. 4; Table S4.1), 203 genes were differentially expressed ($\log_2$ fold change > 2; $P < 0.05$). Among them, 109 genes were upregulated

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

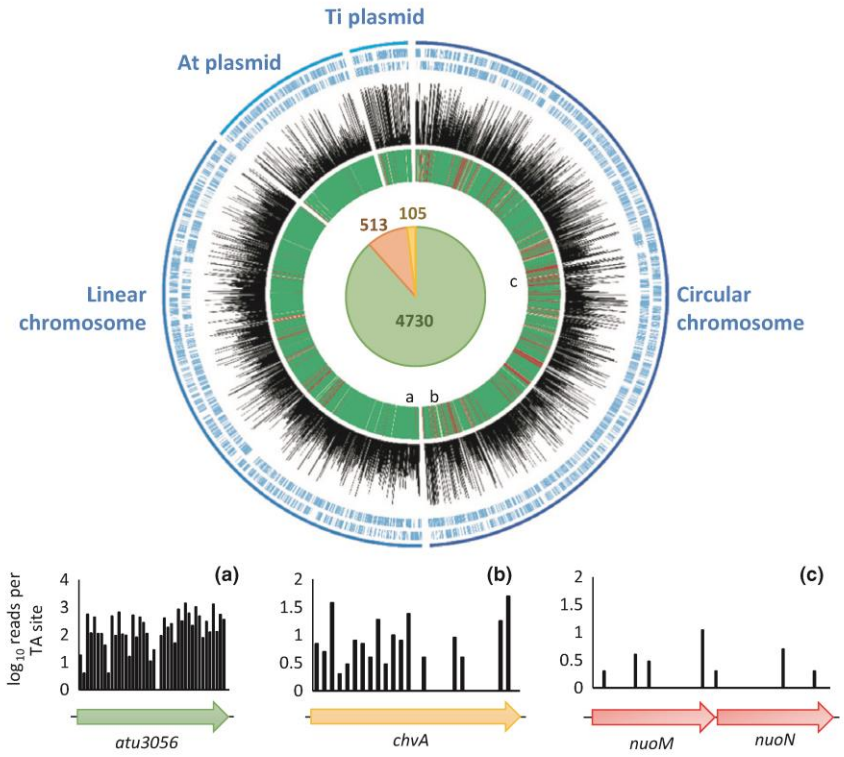


Fig. 2 Genome of *Agrobacterium tumefaciens* C58 with locations of transposon insertions in the constructed transposon library. From the outside to the inside, the tracks represent: forward and reverse coding sequences (in blue), number of transposon insertions per TA site for each gene expressed in $\log_{10}$ (in black), and El-Artist analysis in which nonessential genes are in green, genes with an essential domain are in yellow, and essential genes are in red. The circle chart shows the total number of nonessential genes (in green), essential genes (in red), and genes with a domain essential (in yellow). Panels (a), (b) and (c) exemplify these three categories of genes, showing the number of transposon insertions per TA sites in each of the genes *atu3056*, *chvA*, *nuoM* and *nuoN*.

and 94 were downregulated in the GHB condition. In the GHB condition, the top 10 of the highest upregulated genes ($\log_2$ fold change between 6.44 and 4.17) encompassed the *blcABC* operon and *pckA* gene, which, except for the lactonase-encoding *blcC* gene, were all also identified by Tn-seq as crucial under GHB assimilation. The *blcC* gene encodes a lactonase that is involved in γ-butyrolactone cleavage, but not in GHB degradation (Carlier *et al.*, 2004; Chai *et al.*, 2007). Other remarkable upregulated genes were *sdhCD* (also revealed by Tn-seq), *atu3740* and *pfp* (*atu2115*) encoding two successive steps converting glyceraldehyde 3-phosphate to fructose 6-phosphate in gluconeogenesis, and *dctA* (*atu3298*) coding for a transporter of C_4-dicarboxylic acids. Most of the other upregulated genes belong to the COG category of energy production and conversion, including oxidative phosphorylation pathways (*cyd* and *fix* genes) and nitrate reductase (*nap* genes).

Considering the upregulated genes in the sucrose condition, the most remarkable gene was *kdgA* (*atu4494*) that is coding for the last step of the Entner–Doudoroff pathway. Some others were involved in sugar uptake, such as the *agl* genes (*atu0590–atu0594* coding for a transcriptional regulator, a sugar ATP-binding cassette transporter and a glucosidase) and the *rbs* genes (*atu4369–atu4372* coding a sugar ATP-binding cassette transporter). Most of the other upregulated genes belonged to the COG category inorganic ion transport and metabolism, including iron siderophore synthesis (*atu3670–atu3673* and *atu3675–atu3685*) and uptake (*atu5311–atu5316*) and copper resistance genes (*atu3990–atu3992*). The differentially expressed genes related to C conversion from GHB and sucrose were positioned in a simplified scheme of metabolic pathways (Fig. 5).

When the sucrose–GABA vs sucrose–NH_4 transcriptomes were compared, 163 genes were differentially expressed ($\log_2$ fold change > 2; $P < 0.05$). Most of them (109) were upregulated in the GABA–sucrose condition. In the top five of the highest upregulated genes ($\log_2$ fold change between 5.4 and 4.8) were the aforementioned *blcABC* genes (Fig. 4; Table S4.2). Among the upregulated genes, we searched for putative GABA-transaminase genes coding for the conversion of GABA to SSA. We found two candidate genes, *atu4761* and *atu3407*, highlighting a potential redundancy of this activity in *A. tumefaciens* C58. The gene *atu4761* was co-expressed with *atu4762*, a *blcA* paralogous gene coding for an SSA dehydrogenase. Most of the other upregulated genes belonged to COG category amino acid transport and metabolism, including several transporters (*amtB* = *atu2758*; genes *atu1387–atu1391* and *atu3903–atu3905*) and regulatory proteins (*glnK* = *atu2757*). Remarkably, expression of the genes coding catalase KatE (*atu5491*) and superoxide dismutase SodB (*atu4726*) was also enhanced, indicating a response to an oxidative stress.

Validation of *A. tumefaciens* fitness traits when exploiting the host plant

We constructed single and double mutants of the genes *atu4761* and *atu3407* coding for the putative GABA transaminases. None of these mutants was impaired for growth on GABA as a sole N source (Fig. S4), suggesting that either they are not coding for

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

New Phytologist

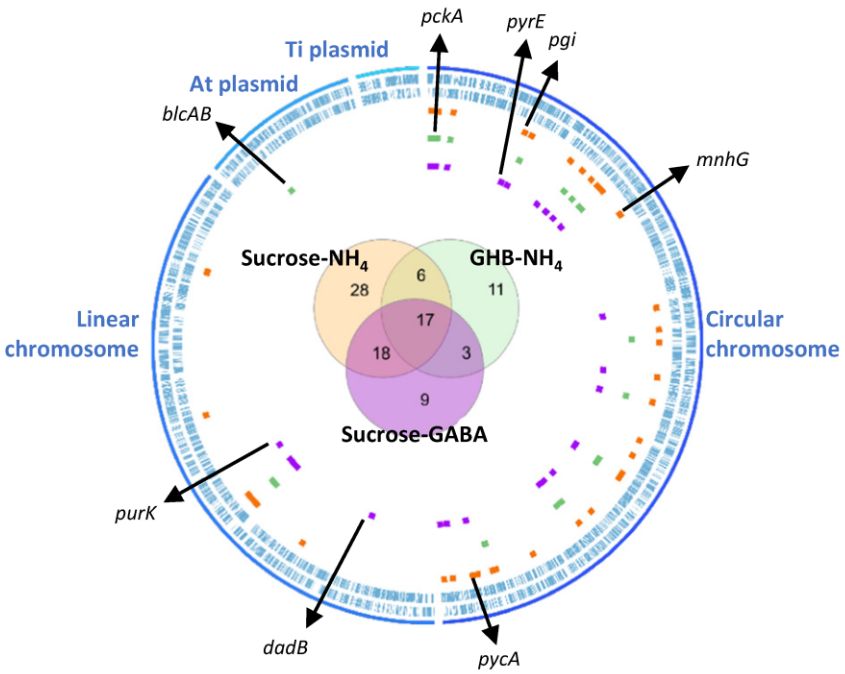


Fig. 3 *Agrobacterium tumefaciens* C58 fitness genes for growth with different carbon and nitrogen sources. From the outside to the inside, the tracks represent: forward and reverse coding sequence (in blue) and fitness genes ('essential genes' according to Con-Artist analysis) when *A. tumefaciens* grew in the presence of sucrose and ammonium (NH_4, as chloride; in orange), γ-hydroxybutyric acid (GHB) and NH_4 (in green), and sucrose and γ-aminobutyric acid (GABA; in purple). The Venn diagram represents a comparison of fitness genes in the three conditions. Some examples of fitness genes are indicated.

GABA transaminase or their mutation was compensated by one (or more) other genes expressing GABA transaminase activity.

We pursued our investigations on C metabolism by comparing GHB and sucrose pathways as archetypes of two major C entries (the TCA cycle and Entner–Doudoroff pathway) in *A. tumefaciens* when it lives on the host plant. Based on the Tn-seq and transcriptomics data, we constructed two *A. tumefaciens* C58 mutants, *pckA::Gm* and *pycA::Gm*, which are affected in pivotal reactions connecting the TCA cycle to gluconeogenesis and Entner–Doudoroff pathways. We also used an already constructed mutant *blcRCAB::Km* deleted for the *blcRCAB* gene cluster (Carlier *et al.*, 2004). Two other *A. tumefaciens* C58 derivatives, 107-Km and 107-Gm, carrying a Km- or Gm-resistance cassette, respectively, in the same noncoding region were used as controls (Haudecoeur *et al.*, 2009a). We verified that the *pycA::Gm* was impaired for growth on sucrose, fructose, and glucose, and the *pckA::Gm* mutant on GHB, succinate, and nopaline as C source. Growth of both mutants was impaired on pyruvate. The mutant *blcRCAB::Km* was only impaired in the GHB assimilation (Fig. 6a). The control strains 107-Km and 107-Gm grew on all C sources.

All these mutants were tested for aggressiveness (tumor weight), proliferation (bacterial charge), and competitive fitness (against wild-type allele) on the tomato host plant. For single strain inoculation experiments, the weight and bacterial charge of 5-wk-old tumors were measured (Fig. 6b,c). The *A. tumefaciens* derivatives 107-Km and 107-Gm exhibited similar traits on the host plant and were used as control conditions. When *blcRCAB::Km* and 107-Km derivatives, which harbor the same Km-resistance cassette, were compared, a decrease of both tumor weight and pathogen charge were observed in the *blcRCAB* mutant. When *pycA::Gm*, *pckA::Gm*, and 107-Gm mutants were compared, a decrease of tumor weight and bacterial charge was observed in the *pycA::Gm* mutant only. A previous study reported a decreased aggressiveness of a *pckA* mutant (Liu *et al.*, 2005), but the virulence assay conditions, and hence resource availability, could explain this discrepancy: stem of entire tomato plants (our study) vs tobacco leaf disks (Liu *et al.*, 2005).

Dual competitions were performed for evaluating the fitness of the *blcRCAB::Km*, *pycA::Gm*, and *pckA::Gm* mutants compared with the control derivatives 107-Km or 107-Gm. All the three mixed populations reached a bacterial charge of 10^6 CFU per tumor (Fig. 7a). The *blcRCAB::Km*, *pycA::Gm*, and *pckA::Gm* mutants were impaired in competitive fitness (Fig. 7b). Finally, we performed competitions between the *pycA::Gm* and *pckA::Gm* mutants to ascertain whether one of the two pathways (Entner–Doudoroff or gluconeogenesis) could be a major contributor to bacterial fitness in plant tumors. The competitive index was close to 1, showing that the two pathways contributed equally to the tumor niche exploitation by *A. tumefaciens*. Noticeably, the mixed population composed of the *pycA::Gm* and *pckA::Gm* mutants colonized the plant tumors less efficiently (bacterial charge in Fig. 7a) compared with the other mixed populations, especially the *pycA::Gm* and C107-Gm mix. This result suggested that *A. tumefaciens* could gain an advantage in the simultaneous expression of the Entner–Doudoroff pathway and gluconeogenesis.

Discussion

The biotrophic pathogen *A. tumefaciens* diverts the plant development and metabolism for constructing and exploiting a privileged ecological niche: the plant tumor. Numerous studies have deciphered the tumor niche construction process by studying the

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

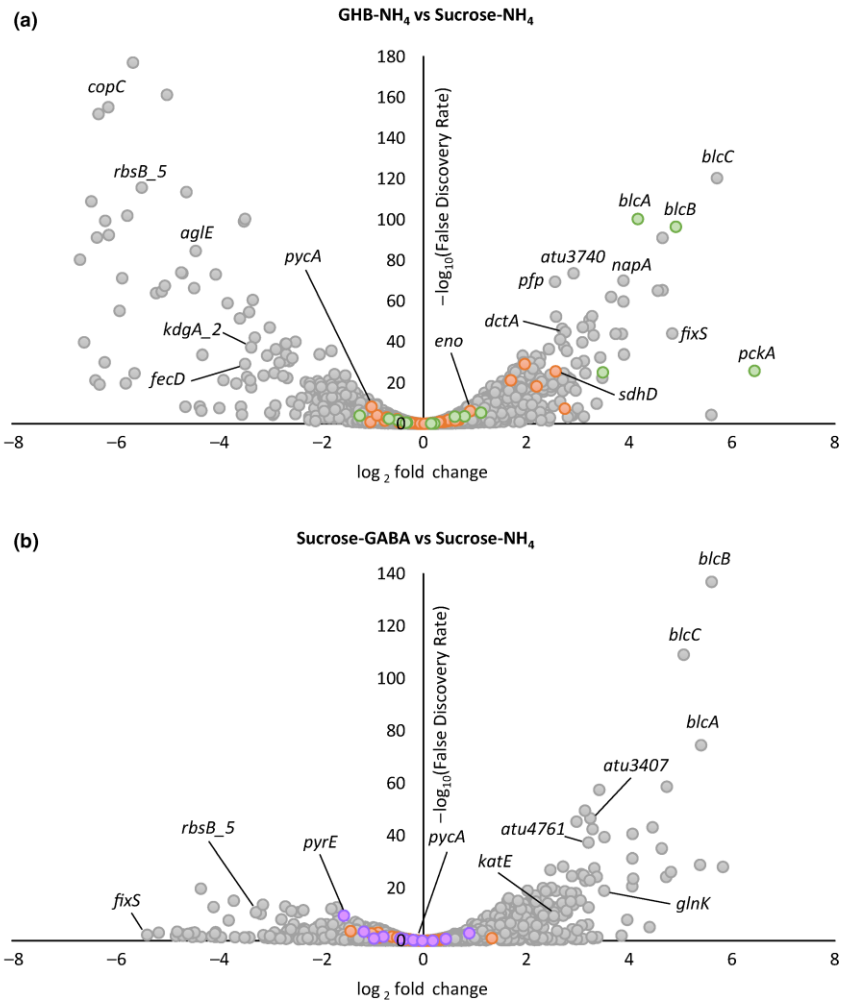


Fig. 4 Volcano plot of transcriptomic data. The data for all genes are plotted as $\log_2$ fold change vs the $-\log_{10}$ of the adjusted *P*-value. (a) Differentially expressed genes between γ-hydroxybutyric acid (GHB)–ammonium (NH_4, as chloride) and sucrose–NH_4 growth conditions. The fitness genes (identified by transposon-sequencing (Tn-seq)) in the GHB–NH_4 growth condition are colored in green, and those in the sucrose–NH_4 growth condition are presented in orange. (b) Differential expressed genes between sucrose–γ-aminobutyric acid (GABA) and sucrose–NH_4 growth conditions. The essential genes (identified by Tn-seq) in the sucrose–GABA condition are in purple, and those in the sucrose–NH_4 condition are presented in orange. The identities of some genes are indicated.

T-DNA transfer and expression in the host plant, as well as mechanisms to escape plant defense (Gohlke & Deeken, 2014; Gelvin, 2017). In this study, we combined different omics (metabolomics, transcriptomics, and Tn-seq) to uncover the role of *A. tumefaciens* genes and pathways in tumor niche exploitation.

Metabolomics of tumor tissues induced on *S. lycopersicum* by *A. tumefaciens* revealed a wide variety of metabolites (e.g. sugars, polyols, amino acids, organic acids, phenolics). They are potential nutrients supporting the proliferation of *A. tumefaciens* that reached 10^6 CFU g^{-1} of fresh tumor tissues. Most of the quantified metabolites (106 of 130) were accumulated at a quite similar concentration in uninfected stems and tumors (Fig. 1). The most abundant metabolites were also the common ones in the two tissues (e.g. glucose, sucrose, malate, glutamate), as well as in tomato seeds and root exudates (Kamilova *et al.*, 2006). These plant metabolites could support the growth of *A. tumefaciens* when it colonizes either asymptomatic or symptomatic plants.

The 24 other metabolites that we quantified, such as GABA, proline, pyruvate, GHB, SSA, and opines, were enriched at least four times in plant tumors compared with healthy stems. Some of them (e.g. agrocinopine, pyruvate, SSA) were increased in plant tumors by several orders of magnitude. These enriched compounds are chemical signatures of the tumor niche: *A. tumefaciens* was expected to have evolved pathways for detoxifying and exploiting these compounds as nutrients and signals. This paradigm is well exemplified by the two opines nopaline and agrocinopine and ferulic derivatives. Nopaline confers a selective growth advantage to a nopaline-assimilating *A. tumefaciens* in *S. lycopersicum* tumors (Lang *et al.*, 2014). The agrocinopine is cleaved into sucrose and arabinose-2-phosphate, which plays an important signaling role: arabinose-2-phosphate enhances the quorum-sensing, which in turn activates the horizontal transfer of the Ti plasmid, and hence the dissemination of the virulence genes (El Sahili *et al.*, 2015). *Agrobacterium* detoxifies ferulic derivatives using different pathways (Brencic *et al.*, 2004; Campillo *et al.*, 2014).

By combining transcriptomics, Tn-seq, and plant infection assays, we investigated the degradative pathways of one common metabolite, sucrose, and two tumor-enriched metabolites, GHB as a C source and GABA as an N source. In the case of sucrose, the combination of Tn-seq and transcriptomics led us to identify the assimilative circuit that starts by the conversion of glucose and fructose into glucose 6-phosphate (*pgi* as a fitness

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

New Phytologist

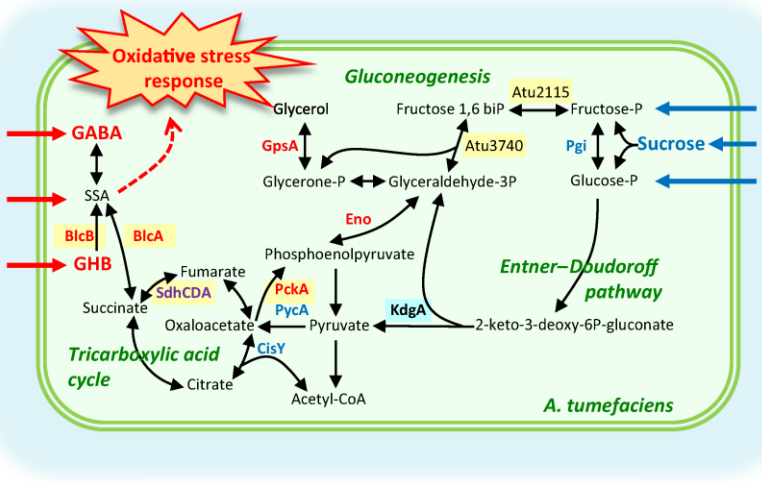


Fig. 5 *Agrobacterium tumefaciens* key pathways for exploiting sucrose and γ-hydroxybutyric acid (GHB). This scheme combines the transposon-sequencing (Tn-seq) and transcriptomics data. Tn-seq revealed fitness genes for growing on either GHB (*blcAB*, *pckA*, *eno*, *gpsA* in red) or sucrose (*pgi*, *pycA*, *cisY* in blue), and on both carbon sources (*sdhCDA* in purple). Transcriptomics revealed response genes when *A. tumefaciens* was grown on either GHB (*blcAB*, *sdhCDA*, *pckA*, *atu3740*, *atu2115* in yellow) or sucrose (*kdgA* in blue). GABA, γ-aminobutyric acid; SSA, succinic semialdehyde.

gene), then its conversion into pyruvate via the Entner–Doudoroff pathway (*kdgA* as an upregulated gene) before entering into the TCA cycle (*pycA*, *cisY*, and *sdhCDA* as fitness genes). These pathways are consistent with a previous metabolic study showing the absence of glycolysis in *A. tumefaciens* and demonstrating the Entner–Doudoroff pathway as a main road of sugar degradation (Fuhrer *et al.*, 2005). The same work also pointed to a high C flux (almost 100%) between the Entner–Doudoroff pathway and the TCA cycle. These two metabolic characteristics are shared by Rhizobiaceae such as *A. tumefaciens* and the legume symbiont *Sinorhizobium meliloti*, and contrasted to sugar assimilation in other bacteria such as *E. coli*, *Bacillus subtilis*, and *Pseudomonas fluorescens* (Fuhrer *et al.*, 2005). The importance of the TCA cycle in a complete exploitation of sugars as resource by *A. tumefaciens* was also supported by our Tn-seq data, with *pycA*, *cisY*, and *sdhCDA* as key fitness genes in the presence of sucrose.

These coherent results led us to evaluate the role of the C flows into and from the TCA cycle in the *A. tumefaciens*–host plant interaction using a reverse genetics approach. In co-infection assays in plant tumors, the *pycA* and *pckA* mutants were outcompeted by a wild-type genotype, but they showed a similar relative fitness when they were competed together. A wild-type *A. tumefaciens* could be considered as a generalist for the assimilation of a wide spectrum of metabolites. By contrast, the constructed *pycA* and *pckA* mutants could be considered as specialists for a restricted range of metabolites that are assimilated by either the gluconeogenesis or the Entner–Doudoroff pathway. A fitness decrease of each mutant in competition with the wild-type highlighted the advantage that the biotrophic pathogen gained by assimilating a wide spectrum of plant metabolites. Moreover, the co-existence of the two *pycA* and *pckA* mutants revealed that the two types of resources were abundant enough and/or differentially distributed to sustain the growth of these two specialists in plant tumors. A remarkable study reported an increase of the sucrose concentration in tumors according to the age of the tumors, as well as in a gradient from the center to the periphery of the tumors on *Ricinus communis* (Wächter *et al.*, 2003). In further studies, the spatial and temporal distribution of the metabolites should be considered as an important parameter driving the resource exploitation strategy of *A. tumefaciens* when it colonizes the heterogeneous environment that the plant tumors are.

Our data also revealed that a mixture of the two specialists (the *pycA* and *pckA* mutants) was less efficient for exploiting the host in terms of bacterial charge than a mixture containing the generalist (the wild-type) and one of the two specialists (Fig. 7). This suggested that *A. tumefaciens* could take advantage of the simultaneous expression of the Entner–Doudoroff pathway and gluconeogenesis in the same individual, quite simply because these two pathways are connected for allowing the recycling of several metabolites such as glyceraldehyde 3-phosphate and phosphoenolpyruvate (Fig. 5). As already discussed, some other explanations related to spatial and temporal distribution of the resource cannot be excluded. The transcriptome of *A. tumefaciens* living in tumors on the *A. thaliana* host plant is consistent with a simultaneous expression of pathways to exploit of a wide diversity of C and N sources (González-Mula *et al.*, 2018). In *A. tumefaciens* and in some other Rhizobiaceae, the separation of C flows incoming and outgoing into and from the TCA cycle using the Entner–Doudoroff pathway and gluconeogenesis (instead of a unique, reversible glycolysis/gluconeogenesis pathway) could be a biological innovation contributing to an optimal exploitation of the diversified resources available in the plant hosts. The capacity of the microbial pathogens to activate and regulate their C assimilative pathways is crucial for survival and invasion in plant and animal hosts (Alteri *et al.*, 2009; Brock, 2009; Basu *et al.*, 2018).

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

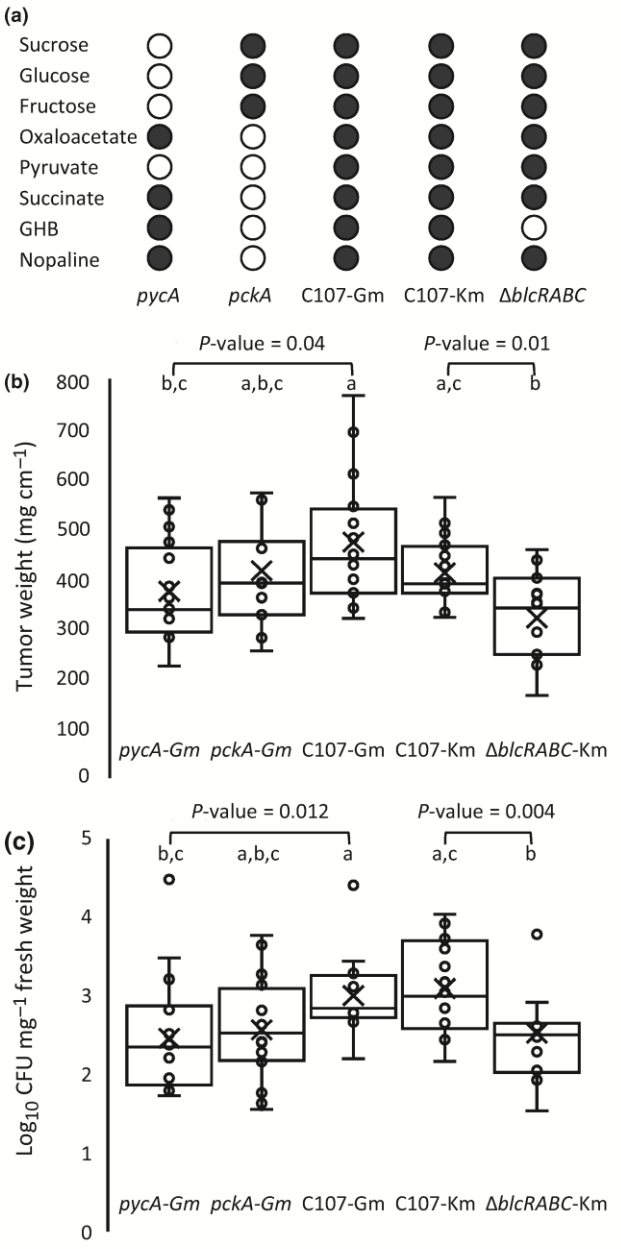


Fig. 6 Metabolic capacity, aggressiveness, and colonization of the constructed *Agrobacterium tumefaciens* mutants. (a) *A. tumefaciens* knockout (KO) mutants and control strains growing in AB medium supplemented with different carbon sources. Open circles represent the absence of growth ($OD_{600} < 0.05$). (b) Fresh weight (FW) of tomato tumors induced with KO mutants and control strains. (c) Colonization efficiency (bacterial numeration, $\log_{10}$ CFU mg^{-1} FW) in tomato tumors. Mean values (indicated by a cross), median (horizontal line), and standard deviations (SD) of two independent experiments are presented. Nonparametric Kruskal–Wallis and post-hoc Tukey tests ($n = 16$; $P < 0.05$) were used, and different letters indicate statistical significance.

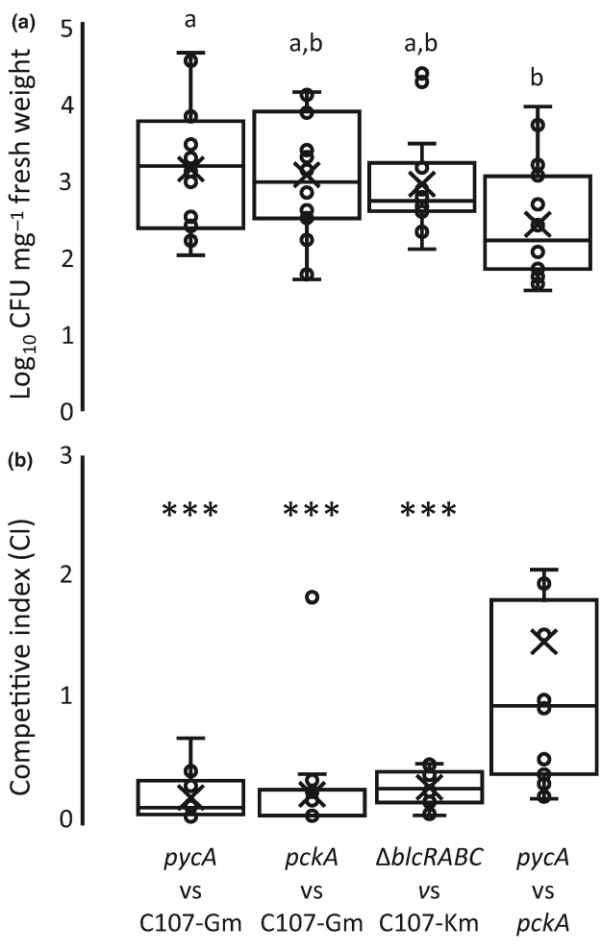


Fig. 7 Competitive fitness of the constructed *Agrobacterium tumefaciens* mutants. (a) Colonization efficiency (total bacterial numeration (mutant plus control), $\log_{10}$ CFU mg^{-1} FW) in tomato tumors. Mean values (indicated by a cross), median (horizontal line), and standard deviations (SD) of two independent experiments are presented. Nonparametric Kruskal–Wallis and post-hoc Tukey tests ($n = 16$; $P < 0.05$) were used and different letters indicate statistical significance. (b) Relative abundance of knockout mutant and control strain was compared at infection time and in tumors. A competitive index value < 1 indicates a fitness loss of the mutant strain in plant tumors. Average (indicated by a cross), median (horizontal line), and SD were calculated from two independent experiments ($n = 16$). Significant fitness loss of mutants is noted by a triple asterisk (Wilcoxon signed rank test $P < 0.001$).

The tumor-enriched metabolites GHB, GABA, and SSA are metabolically connected. GABA is the highest abundant nonprotein amino acid in tumor tissues of *A. thaliana* and *S. lycopersicum* (Deeken *et al.*, 2006; Lang *et al.*, 2016; this work). In the host plant, GABA is mainly produced from glutamate (by GABA decarboxylase) and then degraded into SSA (by GABA transaminase), which is in turn converted into succinate (by SSA dehydrogenase) or GHB (by GHB reductase) (Bown & Shelp, 2016). SSA is a toxic metabolite provoking an oxidative stress in plants and other organisms and microorganisms, including *A. tumefaciens* (Bouché *et al.*, 2003; Ludewig *et al.*, 2008; Wang *et al.*, 2016). In plant tumors, *A. tumefaciens* may exploit plant GABA and GHB as N and C sources but has to face toxic SSA – either exogenous SSA resulting from plant metabolism or

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

New Phytologist

Research 465

endogenous SSA as an intermediate of the *A. tumefaciens* GABA and GHB degradation pathways. The published transcriptome of *A. tumefaciens* in plant tumors showed that the pathogen responded to the presence of GABA and its derived metabolites SSA and GHB, since the *blc* genes, as well as the *atu4761* gene coding for a putative GABA transaminase were upregulated (González-Mula *et al.*, 2018).

When *A. tumefaciens* was grown on GABA as an N source, the Tn-seq approach failed to identify any genes coding for a putative GABA transaminase, nor an SSA dehydrogenase that would be involved in the degradation of GABA and the detoxification of SSA. This may be explained by a redundancy of genes coding these two enzymatic activities. Transcriptomics supported this hypothesis. Two genes coding for putative transaminases (*atu3407* and *atu4761*) and two others for SSA dehydrogenases (*atu4762* and *blcA*) were upregulated in *A. tumefaciens* growing on GABA. The role of the SSA dehydrogenases in stress response and quorum-sensing signal decay was previously studied by Wang *et al.* (2006). In our study, simple and double-KO mutants of the two transaminases still grew on GABA, suggesting the presence of at least a third gene encoding a GABA transaminase in *A. tumefaciens*. In the related species *Rhizobium leguminosarum*, three GABA transaminases are involved in the degradation of GABA (Prell *et al.*, 2009).

When *A. tumefaciens* was grown in the presence of GHB as a C source, a combination of Tn-seq and transcriptomics identified the *blcAB* genes, which are required for the conversion of GHB to succinate (Carlier *et al.*, 2004; Chai *et al.*, 2007). This approach also permitted connecting this particular pathway to the central metabolism by highlighting genes of the TCA cycle, gluconeogenesis, and synthesis of lipid precursors (*sdhDC*, *pckA*, *eno*, *atu3740*, *atu2115* and *gpsA*). The *A. tumefaciens* mutants defective in *blc* or *pckA* genes were unable to grow on GHB as a nutrient, validating the data collected from Tn-seq. Tn-seq and transcriptomics appeared as complementary for deciphering microbial pathways.

Besides a potential growth advantage related to nutrient exploitation, the GABA- and GHB-transcriptomes highlighted an oxidative stress response in *A. tumefaciens*. In the presence of GABA, the upregulated genes concerned were, for instance, *katE* and *sodB*, coding for catalase and superoxide dismutase, respectively. In the presence of GHB, the stress response's upregulated genes were the *cyd* and *fix* genes coding for oxidative phosphorylation pathways, whereas the downregulated genes were involved in siderophore synthesis and uptake (Fig. 4a). A decrease of iron uptake would contribute to reduce the production of highly deleterious hydroxyl radicals via the Fenton reaction. In culture assays, Wang *et al.* (2016) showed that a pre-exposure of *A. tumefaciens* to extracellular SSA induces an oxidative stress response and increases resistance of *A. tumefaciens* to hydrogen peroxide. Noticeably, this effect was lost in a *blcABC* KO-mutant (Wang *et al.*, 2016). Our study showed that a *blc* KO-mutant was impaired for inducing tumors on tomato stems, as well as colonizing plant tumors and competing with a wild-type strain in the plant tumors. Two nonexclusive explanations of this selective advantage could be proposed: an impaired assimilation of GHB or GABA as nutrients, and an impaired SSA-mediated activation of the oxidative stress response to face plant defense.

Different arguments supported the impaired oxidative stress response as an important cause of the decreased aggressiveness and fitness in the *blc* mutant: first, *A. tumefaciens* mutants of catalase and superoxide dismutase were impaired in virulence, highlighting oxidative stress response as an important trait during plant infection (Xu & Pan, 2000; Saenkham *et al.*, 2007); second, whereas the *blc* mutant was affected in aggressiveness (tumor weight), bacterial invasion (bacterial charge), and fitness (competition vs wild-type allele), the *pckA* mutant was impaired in competitive fitness only, suggesting that the *blc* genes conferred an advantage that could not be exclusively explained by nutrition.

In *A. tumefaciens*, the *blc* genes are carried by the dispensable At plasmid, which reaches a size of 0.5 Mb in *A. tumefaciens* C58 (Goodner, 2001; Wood *et al.*, 2001). Several studies have highlighted the fitness cost imposed by maintenance and expression of At plasmid genes (Morton *et al.*, 2013; Platt *et al.*, 2014; González-Mula *et al.*, 2018). This study showed the fitness gains conferred by *blc* genes in plant host infection. The selective advantage conferred by the *blc* operon would not be restricted to *Agrobacterium* pathogens, as data mining analysis revealed its presence in the genome of several host-interacting bacteria, such as *Rhizobium etli*, *Burkholderia phenoliruptrix*, and *Pantoea* sp.

Beyond the use of opines, our study expanded the ecological traits supporting exploitation of tumor niche by *A. tumefaciens*, highlighting novel targets for controlling its virulence and proliferation.

Acknowledgements

We thank Julien Lang, Anthony Kwasiborski and Fabienne Pierre (I2BC) for their contribution to metabolomics, transcriptomics, and reverse genetics, and Erwan Gueguen (MAP, Université Lyon) for kindly providing plasmid pSAM_DGm. This work has benefited from the facilities and expertise of the high-throughput sequencing platform and the plant culture facilities of the I2BC and the metabolomics platform of the Plant Observatory–Chemistry and Metabolomics (Versailles, France). This work was supported by CNRS (SE2016-2017), University Paris-Sud (PhD grant to AG-M) and LabEx Saclay Plant Sciences-SPS (ANR-10-LABX-0040-SPS). The authors declare no conflict of interest.

Author contributions

PM and DF designed research; AG-M, JL, LM and DN performed research; AG-M, DF, JL and FL analyzed data; AG-M, DN, PM and DF wrote the paper.

ORCID

Denis Faure https://orcid.org/0000-0002-5379-8867
Florian Lamouche https://orcid.org/0000-0001-6232-5829
Peter Mergaert https://orcid.org/0000-0002-5919-7317

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

References

Alteri CJ, Smith SN, Mobley HLT. 2009. Fitness of *Escherichia coli* during urinary tract infection requires gluconeogenesis and the TCA cycle. *PLoS Pathogens* **5**: e1000448.

Aznar A, Chen NW, Thomine S, Dellagi A. 2015. Immunity to plant pathogens and iron homeostasis. *Plant Science* **240**: 90–97.

Barton IS, Fuqua C, Platt TG. 2018. Ecological and evolutionary dynamics of a model facultative pathogen: *Agrobacterium* and crown gall disease of plants. *Environmental Microbiology* **20**: 16–29.

Basu P, Sandhu N, Bhatt A, Singh A, Balhana R, Gobe I, Crowhurst NA, Mendum TA, Gao L, Ward JL *et al.* 2018. The anaplerotic node is essential for the intracellular survival of *Mycobacterium tuberculosis*. *Journal of Biological Chemistry* **293**: 5695–5704.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bouché N, Fait A, Bouchez D, Møller SG, Fromm H. 2003. Mitochondrial succinic-semialdehyde dehydrogenase of the gamma-aminobutyrate shunt is required to restrict levels of reactive oxygen intermediates in plants. *Proceedings of the National Academy of Sciences, USA* **100**: 6843–6848.

Bown AW, Shelp BJ. 2016. Plant GABA: not just a metabolite. *Trends in Plant Science* **21**: 811–813.

Brencic A, Eberhard A, Winans SC. 2004. Signal quenching, detoxification and mineralization of *vir* gene-inducing phenolics by the VirH2 protein of *Agrobacterium tumefaciens*. *Molecular Microbiology* **51**: 1103–1115.

Brock M. 2009. Fungal metabolism in host niches. *Current Opinion in Microbiology* **12**: 371–376.

Campillo T, Renoud S, Kerzaon I, Vial L, Baude J, Gaillard V, Bellvert F, Chamignon C, Comte G, Nesme X *et al.* 2014. Analysis of hydroxycinnamic acid degradation in *Agrobacterium fabrum* reveals a coenzyme A-dependent, beta-oxidative deacetylation pathway. *Applied and Environmental Microbiology* **80**: 3341–3349.

Carlier A, Chevrot R, Dessaux Y, Faure D. 2004. The assimilation of γ-butyrolactone in *Agrobacterium tumefaciens* C58 interferes with the accumulation of the *N*-acyl-homoserine lactone signal. *Molecular Plant–Microbe Interactions* **17**: 951–957.

Chai Y, Ching ST, Cho H, Winans SC. 2007. Reconstitution of the biochemical activities of the AttJ repressor and the AttK, AttL, and AttM catabolic enzymes of *Agrobacterium tumefaciens*. *Journal of Bacteriology* **189**: 3674–3679.

Chevrot R, Rosen R, Haudecoeur E, Cirou A, Shelp BJ, Ron E, Faure D. 2006. GABA controls the level of quorum-sensing signal in *Agrobacterium tumefaciens*. *Proceedings of the National Academy of Sciences, USA* **103**: 7460–7464.

Chilton MD, Currier TC, Farrand SK, Bendich AJ, Gordon MP, Nester EW. 1974. *Agrobacterium tumefaciens* DNA and PS8 bacteriophage DNA not detected in crown gall tumors. *Proceedings of the National Academy of Sciences, USA* **71**: 3672–3676.

Christen B, Abeliuk E, Collier JM, Kalogeraki VS, Passarelli B, Coller JA, Fero MJ, McAdams HH, Shapiro L. 2011. The essential genome of a bacterium. *Molecular Systems Biology* **7**: E528.

Curtis PD, Brun YV. 2014. Identification of essential alphaproteobacterial genes reveals operational variability in conserved developmental and cell cycle systems. *Molecular Microbiology* **93**: 713–735.

Deeken R, Engelmann JC, Efetova M, Czirjak T, Muller T, Kaiser WM, Tietz O, Krischke M, Mueller MJ, Palme K *et al.* 2006. An integrated view of gene expression and solute profiles of *Arabidopsis* tumors: a genome-wide approach. *Plant Cell* **18**: 3617–3634.

DeJesus MA, Ioerger TR. 2013. A hidden Markov model for identifying essential and growth-defect regions in bacterial genomes from transposon insertion sequencing data. *BMC Bioinformatics* **14**: e303.

Dennis JJ, Zylstra GJ. 1998. Plasposons: modular self-cloning minitransposon derivatives for rapid genetic analysis of Gram-negative bacterial genomes. *Applied and Environmental Microbiology* **64**: 2710–2715.

Dessaux Y, Faure D. 2018. Quorum sensing and quorum quenching in *Agrobacterium*: a 'go/no go system'? *Genes* **9**: E210.

Dessaux Y, Petit A, Tempe J. 1993. Chemistry and biochemistry of opines, chemical mediators of parasitism. *Phytochemistry* **34**: 31–38.

El Sahili A, Li S-Z, Lang J, Virus C, Planamente S, Ahmar M, Guimaraes BG, Aumont-Nicaise M, Vigouroux A, Soulère L *et al.* 2015. A pyranose-2-phosphate motif is responsible for both antibiotic import and quorum-sensing regulation in *Agrobacterium tumefaciens*. *PLoS Pathogens* **11**: e1005071.

Faure D, Simon JC, Heulin T. 2018. Holobiont: a conceptual framework to explore the eco-evolutionary and functional implications of host–microbiota interactions in all ecosystems. *New Phytologist* **218**: 1321–1324.

Fields JH, Baldwin J, Hochachka PW. 1976. On the role of octopine dehydrogenase in cephalopod mantle muscle metabolism. *Canadian Journal of Zoology* **54**: 871–878.

Fuhrer T, Fischer E, Sauer U. 2005. Experimental identification and quantification of glucose metabolism in seven bacterial species. *Society* **187**: 1581–1590.

Gelvin SB. 2017. Integration of *Agrobacterium* T-DNA into the plant genome. *Annual Review of Genetics* **82**: 5542–5552.

Gohlke J, Deeken R. 2014. Plant responses to *Agrobacterium tumefaciens* and crown gall development. *Frontiers in Plant Science* **5**: 155.

González-Mula A, Lang J, Grandclément C, Naquin D, Ahmar M, Soulère L, Queneau Y, Dessaux Y, Faure D. 2018. Lifestyle of the biotroph *Agrobacterium tumefaciens* in the ecological niche constructed on its host plant. *New Phytologist* **219**: 350–362.

Goodner B. 2001. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* **294**: 2323–2328.

Haudecoeur E, Planamente S, Cirou A, Tannières M, Shelp BJ, Moréra S, Faure D. 2009b. Proline antagonizes GABA-induced quenching of quorum-sensing in *Agrobacterium tumefaciens*. *Proceedings of the National Academy of Sciences, USA* **106**: 14587–14592.

Haudecoeur E, Tannieres M, Cirou A, Raffoux A, Dessaux Y, Faure D. 2009a. Different regulation and roles of lactonases AiiB and AttM in *Agrobacterium tumefaciens* C58. *Molecular Plant–Microbe Interactions* **22**: 529–537.

Kamilova F, Kravchenko LV, Shaposhnikov AI, Azarova T, Makarova N, Lugtenberg B. 2006. Organic acids, sugars, and L-tryptophane in exudates of vegetables growing on stonewool and their effects on activities of rhizosphere bacteria. *Molecular Plant–Microbe Interactions* **19**: 250–256.

Kylafis G, Loreau M. 2011. Niche construction in the light of niche theory. *Ecology Letters* **14**: 82–90.

Lang J, Gonzalez-Mula A, Taconnat L, Clement G, Faure D. 2016. The plant GABA signaling downregulates horizontal transfer of the *Agrobacterium tumefaciens* virulence plasmid. *New Phytologist* **210**: 974–983.

Lang J, Vigouroux A, El Sahili A, Kwasiborski A, Aumont-Nicaise M, Dessaux Y, Shykoff JA, Moréra S, Faure D, El Sahili A *et al.* 2017. Fitness costs restrict niche expansion by generalist niche-constructing pathogens. *The ISME Journal* **11**: 374–385.

Lang J, Vigouroux A, Planamente S, El Sahili A, Blin P, Aumont-Nicaise M, Dessaux Y, Moréra S, Faure D. 2014. *Agrobacterium* uses a unique ligand-binding mode for trapping opines and acquiring a competitive advantage in the niche construction on plant host. *PLoS Pathogens* **10**: e1004444.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**: R25.

Liao Y, Smyth GK, Shi W. 2014. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.

Liu P, Wood D, Nester EW. 2005. Phosphoenolpyruvate carboxykinase is an acid-induced, chromosomally encoded virulence factor in *Agrobacterium tumefaciens*. *Journal of Bacteriology* **187**: 6039–6045.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* **25**: 402–408.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: e550.

Ludewig F, Hüser A, Fromm H, Beauclair L, Bouché N. 2008. Mutants of GABA transaminase (POP2) suppress the severe phenotype of succinic semialdehyde dehydrogenase (ssadh) mutants in *Arabidopsis*. *PLoS ONE* **3**: e3383.

Macho AP, Guidot A, Barberis P, Beuzón CR, Genin S. 2010. A competitive index assay identifies several *Ralstonia solanacearum* type III effector mutant strains with reduced fitness in host plants. *Molecular Plant–Microbe Interactions* **23**: 1197–1205.

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

Martin FM, Uroz S, Barker DG. 2017. Ancestral alliances: plant mutualistic symbioses with fungi and bacteria. *Science* **356**: eaad4501.

Marty L, Vigouroux A, Aumont-Nicaise M, Dessaux Y, Faure D, Moréra S. 2016. Structural basis for high specificity of Amadori compound and mannopine opine binding in bacterial pathogens. *Journal of Biological Chemistry* **291**: 22638–22649.

McNally L, Brown SP. 2015. Building the microbiome in health and disease: niche construction and social conflict in bacteria. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **370**: e20140298.

Morton ER, Merritt PM, Bever JD, Fuqua C. 2013. Large deletions in the pAtC58 megaplasmid of *Agrobacterium tumefaciens* can confer reduced carriage cost and increased expression of virulence genes. *Genome Biology and Evolution* **5**: 1453–1464.

van Opijnen T, Camilli A. 2013. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nature Reviews Microbiology* **11**: 435–442.

Planamente S, Vigouroux A, Mondy S, Nicaise M, Faure D, Moréra S. 2010. A conserved mechanism of GABA binding and antagonism is revealed by structure–function analysis of the periplasmic binding protein Atu2422 in *Agrobacterium tumefaciens*. *The Journal of Biological Chemistry* **285**: 30294–30303.

Platt TG, Morton ER, Barton IS, Bever JD, Fuqua C. 2014. Ecological dynamics and complex interactions of *Agrobacterium* megaplasmids. *Frontiers in Plant Science* **5**: e635.

Podlešáková K, Ugena L, Spíchal L, Doležal K, De Diego N. 2019. Phytohormones and polyamines regulate plant stress responses by altering GABA pathway. *New Biotechnology* **48**: 53–65.

Poole P, Ramachandran V, Terpolilli J. 2018. Rhizobia: from saprophytes to endosymbionts. *Nature Reviews Microbiology* **16**: 291–303.

Prell J, Bourdés A, Karunakaran R, Lopez-Gomez M, Poole P. 2009. Pathway of γ-aminobutyrate metabolism in *Rhizobium leguminosarum* 3841 and its role in symbiosis. *Journal of Bacteriology* **191**: 2177–2186.

Pritchard JR, Chao MC, Abel S, Davis BM, Baranowski C, Zhang YJ, Rubin EJ, Waldor MK. 2014. ARTIST: high-resolution genome-wide assessment of fitness using transposon-insertion sequencing. *PLoS Genetics* **10**: e1004782.

Saenkham P, Eiamphungporn W, Farrand SK, Vattanaviboon P, Mongkolsuk S. 2007. Multiple superoxide dismutases in *Agrobacterium tumefaciens*: functional analysis, gene regulation, and influence on tumorigenesis. *Journal of Bacteriology* **189**: 8807–8817.

Simoh S, Quintana N, Kim HK, Choi YH, Verpoorte R. 2009. Metabolic changes in *Agrobacterium tumefaciens*-infected *Brassica rapa*. *Journal of Plant Physiology* **166**: 1005–1014.

Skurnik D, Roux D, Aschard H, Cattoir V, Yoder-Himes D, Lory S, Pier GB. 2013. A comprehensive analysis of in vitro and in vivo genetic fitness of *Pseudomonas aeruginosa* using high-throughput sequencing of transposon libraries. *PLoS Pathogens* **9**: e1003582.

Tannières M, Lang J, Barnier C, Shykoff JA, Faure D. 2017. Quorum-quenching limits quorum-sensing exploitation by signal-negative invaders. *Scientific Reports* **7**: 40126.

Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research* **28**: 33–36.

Vigouroux A, El Sahili A, Lang J, Aumont-Nicaise M, Dessaux Y, Faure D, Moréra S. 2017. Structural basis for high specificity of octopine binding in the plant pathogen *Agrobacterium tumefaciens*. *Scientific Reports* **7**: e18033.

Wächter R, Langhans M, Aloni R, Götz S, Weilmünster A, Koops A, Temguia L, Mistrik I, Pavlovkin J, Rascher U *et al.* 2003. Vascularization, high-volume solution flow, and localized roles for enzymes of sucrose metabolism during tumorigenesis by *Agrobacterium tumefaciens*. *Plant Physiology* **133**: 1024–1037.

Wang C, Tang D, Gao YG, Zhang LH. 2016. Succinic semialdehyde promotes prosurvival capability of *Agrobacterium tumefaciens*. *Journal of Bacteriology* **198**: 930–940.

Wang C, Zhang HB, Wang LH, Zhang LH. 2006. Succinic semialdehyde couples stress response to quorum-sensing signal decay in *Agrobacterium tumefaciens*. *Molecular Microbiology* **62**: 45–56.

Wood DW, Setubal JC, Kaul R, Monks DE, Kitajima JP, Okura VK, Zhou Y, Chen L, Wood GE, Almeida NF *et al.* 2001. The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**: 2317–2323.

Xu XQ, Pan SQ. 2000. An *Agrobacterium* catalase is a virulence factor involved in tumorigenesis. *Molecular Microbiology* **35**: 407–414.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 The *A. tumefaciens* C58 essential genes identified in the Tn*5* and *Himar1* transposon libraries.

Fig. S2 Classification of essential genes by COG categories.

Fig. S3 RT-qPCR and transcriptome comparative gene expression.

Fig. S4 Growth on GABA as a sole nitrogen source.

Table S1 Primer list.

Table S2 Plant metabolomics.

Table S3 Tn-seq data.

Table S4 Transcriptomics.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

See also the Commentary on this article by Woodcock & Malone, **222**: 8–10.

© 2018 The Authors
New Phytologist © 2018 New Phytologist Trust

Acceptance date: 3 November 2018

The biotroph *Agrobacterium tumefaciens* thrives in tumors by exploiting a wide spectrum of plant host metabolites

Subtitle: Plant host exploitation by the *Agrobacterium* biotroph

Almudena Gonzalez-Mula, Joy Lachat, Léo Mathias, Delphine Naquin, Florian Lamouche, Peter Mergaert, Denis Faure

Institute for integrative biology of the cell (I2BC), CNRS CEA Univ. Paris-Sud, University Paris-Saclay, Gif-sur-Yvette F-91190, France

Author for correspondence: Denis Faure

Tel: (+33) 169823498

Email: denis.faure@i2bc.paris-saclay.fr

ORCID-0000-0002-5379-8867

Supplementary files:

- **Figure S1. The *A. tumefaciens* C58 essential genes identified in the Tn*5* and *Himar1* transposon libraries.**
- **Figure S2. Classification of essential genes by COG categories.**
- **Figure S3. RT-qPCR and transcriptome comparative gene expression.**
- **Figure S4. Growth on GABA as a sole nitrogen source.**
- **Table S1. Primer list (see separate Excel file).**
- **Table S2. Plant metabolomics (see separate Excel file).**
- **Table S3. Tn-seq data (see separate Excel file).**
- **Table S4. Transcriptomics (see separate Excel file).**

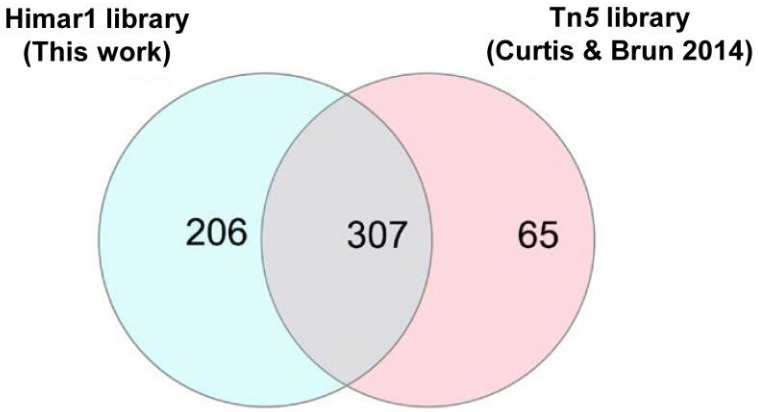


Figure S1. The *A. tumefaciens* C58 essential genes identified in the Tn*5* and *Himar1* transposon libraries.

The Venn diagram shows the essential genes obtained by the Tn*5* library (Curtis & Brun, 2014) and the Himar1 transposon library (this work). The complete list of the genes is available in Table S3.

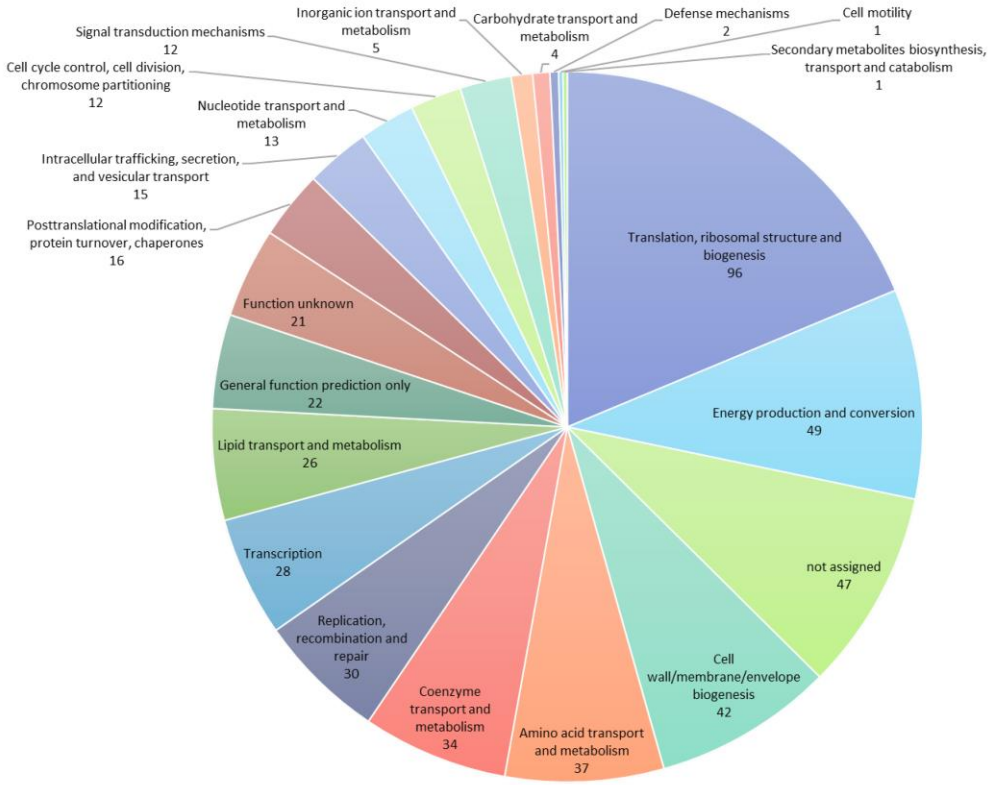


Figure S2. Classification of essential genes by COG categories.

Each number correspond to the total of essential genes of each category in *A. tumefaciens* C58 genome.

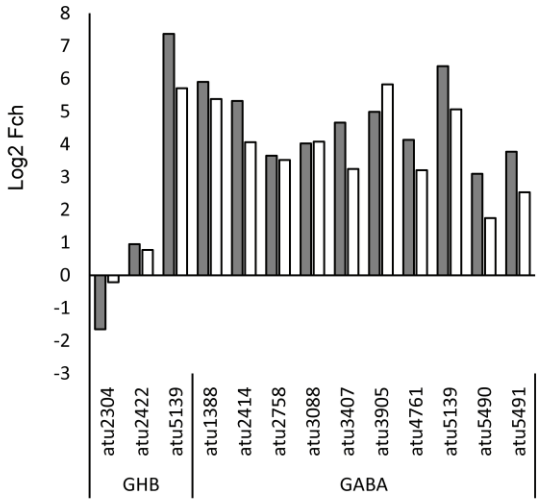


Figure S3. RT-qPCR and transcriptome comparative gene expression.

Comparison of the expression fold change (Log2 Fch) in *A. tumefaciens* C58 cells in GHB-NH_4 vs sucrose-NH_4 and sucrose-GABA vs sucrose-NH_4 growth conditions using quantitative RT-PCR (RT-qPCR in grey) or RNAseq technique (transcriptome in white).

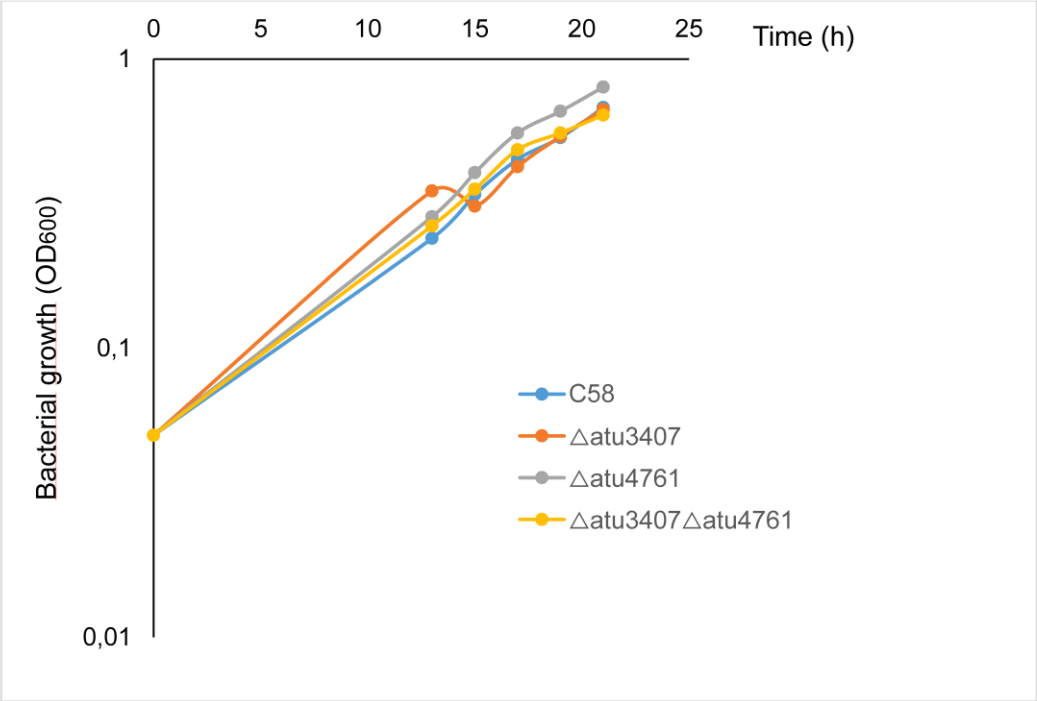


Figure S4. Growth on GABA as a sole nitrogen source.

Growth of the *A. tumefaciens* C58 wild-type and its derivatives (the single KO-mutants *Δatu3407* and *Δatu4761* and the *Δatu3407Δatu4761* double KO-mutant) were similar in AB medium supplemented with sucrose and GABA.

Reference

Curtis PD, Brun Y V. 2014. Identification of essential alphaproteobacterial genes reveals operational variability in conserved developmental and cell cycle systems. Molecular Microbiology 93: 713–735.

4. Comparative cytology, physiology and transcriptomics of free-living and symbiotic *Burkholderia* reveal features for colonization of its insect host

Tsubasa Ohbayashia, Ryo Futahashi, Mia Terashima, Quentin Barriere, Florian Lamouche, Xian-Ying Meng, Yasuo Mitani, Teruo Sone, Shuji Shigenobu, Takema Fukatsu, Peter Mergaert et Yoshitomo Kikuchi

Publié dans la revue *The ISME Journal* (2019). DOI: https://doi.org/10.1038/s41396-019-0361-8

The ISME Journal
https://doi.org/10.1038/s41396-019-0361-8

ARTICLE

Comparative cytology, physiology and transcriptomics of *Burkholderia insecticola* in symbiosis with the bean bug *Riptortus pedestris* and in culture

Tsubasa Ohbayashi[1,2] · Ryo Futahashi[3] · Mia Terashima[2,4] · Quentin Barrière[1] · Florian Lamouche[1] · Kazutaka Takeshita[2,8] · Xian-Ying Meng[3] · Yasuo Mitani[3] · Teruo Sone[2] · Shuji Shigenobu[5] · Takema Fukatsu[3] · Peter Mergaert[1] · Yoshitomo Kikuchi[2,6,7]

Received: 13 April 2018 / Revised: 4 December 2018 / Accepted: 19 January 2019
© International Society for Microbial Ecology 2019

Abstract

In the symbiosis of the bean bug *Riptortus pedestris* with *Burkholderia insecticola*, the bacteria occupy an exclusive niche in the insect midgut and favor insect development and reproduction. In order to understand how the symbiotic bacteria stably colonize the midgut crypts and which services they provide to the host, we compared the cytology, physiology, and transcriptomics of free-living and midgut-colonizing *B. insecticola*. The analyses revealed that midgut-colonizing bacteria were smaller in size and had lower DNA content, they had increased stress sensitivity, lost motility, and an altered cell surface. Transcriptomics revealed what kinds of nutrients are provided by the bean bug to the *Burkholderia* symbiont. Transporters and metabolic pathways of diverse sugars such as rhamnose and ribose, and sulfur compounds like sulfate and taurine were upregulated in the midgut-colonizing symbionts. Moreover, pathways enabling the assimilation of insect nitrogen wastes, i.e. allantoin and urea, were also upregulated. The data further suggested that the midgut-colonizing symbionts produced all essential amino acids and B vitamins, some of which are scarce in the soybean food of the host insect. Together, these findings suggest that the *Burkholderia* symbiont is fed with specific nutrients and also recycles host metabolic wastes in the insect gut, and in return, the bacterial symbiont provides the host with essential nutrients limited in the insect food, contributing to the rapid growth and enhanced reproduction of the bean bug host.

Introduction

Mutualistic bacteria cooperate with their host by providing services, and at the same time, they have to be able to stably colonize the internal environment of the eukaryotic host and to overcome host control mechanisms. To accomplish this, mutualists often alter their morphology, physiology, cytology, and sometimes even their genomic composition [1, 2]. Insects, that feed on nutritionaly unbalanced or undigestable diets, commonly possess mutualistic bacteria in specialized organs, such as gut crypts and bacteriomes [3–5], wherein symbionts play pivotal metabolic roles, including provision

Supplementary information The online version of this article (https://doi.org/10.1038/s41396-019-0361-8) contains supplementary material, which is available to authorized users.

✉ Peter Mergaert
peter.mergaert@i2bc.paris-saclay.fr

✉ Yoshitomo Kikuchi
y-kikuchi@aist.go.jp

[1] Institute for Integrative Biology of the Cell, UMR9198, CNRS, Université Paris-Sud, CEA, Gif-sur-Yvette, France

[2] Graduate School of Agriculture, Hokkaido University, Sapporo, Japan

[3] Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

[4] Institute of Low Temperature Science, Hokkaido University, Sapporo, Japan

[5] NIBB Core Research Facilities, National Institute for Basic Biology, Okazaki, Japan

[6] Computational Bio Big Data Open Innovation Laboratory (CBBD-OIL), AIST, Sapporo, Japan

[7] Bioproduction Research Institute, AIST, Sapporo, Japan

[8] Present address: Faculty of Bioresource Sciences, Akita Prefectural University, Akita, Japan

Published online: 11 February 2019

of essential nutrients [6, 7], digestion of plant cell-wall components [8–10], and reutilization of metabolic wastes of the host [6, 11]. In addition to such nutritional contributions, microbial symbionts are also involved in defense against natural enemies [12, 13], sex determination [14], and detoxification of phytotoxins and pesticides [15]. Most of the insect symbionts are vertically transmitted and highly adapted to their host's internal environment [5, 16]. As a result of, in the evolutionary scale, long-term exclusive residence within the insect hosts, the symbionts tend to lose genes that are not required for life in these unchanging conditions, including genes for flagella motility, cell-wall biosynthesis, transporters, and secretion systems [7, 17, 18]. Such reduced genomes of symbionts are thought to contribute to rapid reproduction and efficient service-provision inside insect hosts.

The bean bug *Riptortus pedestris* (Hemiptera: Alydidae) is a notorious pest of leguminous crops and is broadly distributed in Southeast Asia [19–22]. The insect interacts intimately with a specific symbiont called *Burkholderia insecticola* [23], a dense population of which is harbored in crypts in the posterior midgut region, called M4 (midgut 4th section) [24]. In comparison with symbiotic insects, aposymbiotic insects have a prolonged developmental period, smaller body size and reduced egg numbers [25–27], indicating a pivotal metabolic role of the *Burkholderia* symbiont in the bean bug host. Exceptional among insects, *R. pedestris* and allied species do not vertically transmit their symbionts but instead acquire the specific *Burkholderia* symbiont from the soil every generation [25, 28, 29]. In the bean bug, the symbiont colonization starts and fully establishes within the 2nd instar stage [26, 30]. Since the *Burkholderia* symbiont can be cultivated and genetically manipulated, the *Riptortus–Burkholderia* symbiotic system is an attractive model for clarifying the genetic and molecular basis of insect-microbe symbiotic associations [24, 31]. Previous studies have revealed that the *Burkholderia* symbiont requires its flagellar motility to reach the symbiotic organ [32, 33], and a specific lipopolysaccharide (LPS) structure and other envelope molecules as well as polyhydroxyalkanoate (PHA) metabolism for a full and lasting colonization of the symbiotic organ [34–38].

In *R. pedestris*, the M4 region as well as a distinct, adjacent midgut region called the M4 bulb (M4B), produce specifically a diverse set of digestive enzymes and antimicrobial peptides, including a large family of cysteine-rich peptides which we call the crypt-specific cysteine-rich peptides (CCRs) [36, 39–42]. This suggests that the *Burkholderia* symbiont is exposed to various stresses in the symbiotic organ. It remains largely unknown how the *Burkholderia* symbiont alters its physiology to cope with the specific conditions of the gut environment of the bean bug host. In addition, the metabolic basis of the benefits that the insect host obtains from the bacterial symbionts is still unclear.

To understand the genetic and physiological features and the metabolic activity of *B. insecticola* in the host internal environment, we compared here the cell morphology and cytology by microscopic observations, the bacterial physiology by stress response tests, and transcriptomics by RNA-sequencing (RNA-seq) between the free-living *Burkholderia* symbiont cultured in liquid media (in vitro) and symbiotic *Burkholderia* colonizing the midgut crypts (in vivo).

Materials and methods

Insect and bacterial strains

The bean bug was maintained in the laboratory as described previously [32] using soybeans as food source and distilled water containing 0.05% ascorbic acid as drinking water. *B. insecticola* strain RPE75 [25], a spontaneous rifampicin-resistant mutant (Rif^r) of the wild-type strain RPE64 [23], and its GFP (green fluorescent protein) labeled derivative RPE225 [26] were used in this study.

Preparation of in vitro symbiont cells

The *Burkholderia* symbiont was grown in YG and/or MM medium (Supplementary Table S1) containing rifampicin (10 μg/ml) on a gyratory shaker (150 rpm) at 27 °C, and cultured cells were centrifuged and washed twice with phosphate-buffered saline (PBS) before analysis (Supplementary Table S1).

Preparation of in vivo symbiont cells

The oral administration of the symbiont to the insect and its dissection were performed as described previously [32]. Three days old (second instar) insects were infected by adding the symbiont to their drinking water (10^7 cfu/mL). Symbiotic organs were collected by dissection and the pooled midgut crypts of 10–30 insects were homogenized by a pestle, centrifuged, washed twice with PBS, and then subjected for further analyses. The harvests were performed at the 3rd or 5th instar (5 or 12 days after infection, respectively), depending on the experiments (see Supplementary Table S1).

Stress-sensitivity assays

Suspensions of the in vivo and in vitro symbiont cells were exposed to 13 different antimicrobial chemicals (Supplementary Table S2) for 1 h at 27 °C. These cells were diluted,

inoculated on YG-Rf plates, and incubated at 27 °C for 2 days. Bacterial growth was determined by CFU counting and data are presented as relative CFU numbers. Further details of the method are provided in Supplementary Information.

RNA-sequencing analysis

Total RNA was extracted from triplicate samples from in vivo symbiont cells and 3 h- (mean $OD_{600} = 0.27$), 8 h- (mean $OD_{600} = 1.87$), and 16 h- (mean $OD_{600} = 6.38$) cultured in vitro symbiont cells by using trizol (Thermo Scientific, Wilmington, NC, USA), the RNeasy mini kit (Qiagen Inc., Valencia, CA, USA), and the TURBO DNA-free kit (Thermo Scientific, Wilmington, NC, USA). The mRNA was further enriched by the Eukaryote Human/Mouse/Rat and/or the Bacteria Ribo-Zero kits (Illumina, Inc., San Diego, CA, USA). The cDNA libraries were constructed from approximately 1 µg of total RNA by use of the Scriptseq v2 RNA-seq library kit (Illumina, Inc., San Diego, CA, USA). In total, 12 cDNA libraries were constructed, and sequenced by Nextseq (ver. NS500446, Illumina, Inc., San Diego, CA, USA). The obtained RNA-seq data were analyzed by the bcl2fastq software (ver. 2-2.18.12, Illumina, Inc., San Diego, CA, USA), Cutadapt (ver. 1.15 [43]) for adapter trimming, FastQC (ver. 0.11.5 [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]) for quality control, and BWA (ver. 0.6.2-r126 [44]) for mapping on the symbiont genome (GCA_00402035.1 [45]). Comparisons of gene expression levels were performed by Wald-tests in the DESeq2 program (version 1.12.4 [46]). Genes specifically upregulated and downregulated in vivo were defined as genes whose false discovery rate (FDR) *q*-values were <0.01 and expression fold changes were >2 or <–0.5 when their expressions were compared to 3 h-, 8 h-, and 16 h-cultured in vitro cells. Principal component analysis and heatmap presentations were performed by the FactoMineR (ver. 1.33 [47]) and the gplots package (http://cran.r-project.org/web/packages/gplots/index.html), respectively.

Nucleotide sequence accession number

The RNA-seq nucleotide sequences reported in this study were deposited in the DDBJ/Genbank/EBI databases under the accession numbers: PRJDB6699, DRA007366, and SAMD00140959-SAMD00140970 (Supplementary Table S4).

Other methods

Methods for fluorescence microscopy, flow cytometry, electron microscopy, motility test, quantitative reverse-transcription PCR (qRT-PCR), colony PCR and statistical analyses are described in Supplementary Information.

Results

Morphological and cytological alterations of in vivo *Burkholderia* symbionts

Observed by light microscopy, in vitro symbiont cells were found to be large rods, as expected for a species of the genus *Burkholderia*. In contrast, in vivo symbiont cells appeared as small rods, close to cocci (Fig. 1a–d). The results of flow cytometry analysis by the light scatter corresponded to these microscopy observations and revealed a morphological alteration in the in vivo cells (Fig. 1e). The DAPI staining showed that the in vivo cells had a lower DNA content than the in vitro cells (Fig. 1f). The smaller size of the in vivo symbiont was furthermore obvious by their lower fluorescence intensity after staining with the cell membrane-specific, styryl-based dye FM4-64, (Fig. 1c, h). Together, these data imply that the cell size, DNA content and cell cycle of the *Burkholderia* symbiont undergo a marked shift during the colonization of the symbiotic organ. It may be relevant that in vitro, *Burkholderia* also become small with low DNA content when entering stationary phase or when grown in a minimal medium (see Supplementary Information; Supplementary Fig. S1). Nile red staining demonstrated that PHA (a bacterial carbon storage compound) accumulated in in vivo symbiont cells (Fig. 1b), although the flow cytometry analysis of a large number of cells showed that the PHA accumulation is in the total bacterial population not very different between the in vivo and in vitro bacteria (Fig. 1g).

The observation of in vitro and in vivo cells by TEM corroborated the features of cell shape and PHA accumulation observed by light microscopy and flow cytometry (Fig. 1j, k). Interestingly, frequent membrane blebs in 5 out of 12 cells were found on the surface of in vivo symbiont cells while they were not observed in in vitro cells (Fig. 1k; Supplementary Fig. S2).

The *Burkholderia* symbiont needs flagellar motility to reach and colonize the midgut crypts [32, 48]. However, testing the motility of in vivo and in vitro symbionts revealed that in vivo cells showed little motility at 16 h after incubation (Fig. 1l) but that their flagellar mortility was recovered after 24 h incubation (Supplementary Fig. S3). This suggests that the symbiont loses its motility after having colonized M4 but that this process is reversible and motility is reactivated when the bacteria are released from the symbiotic organ. Taken together, these results demonstrate that the cellular characteristics of the *Burkholderia* are altered in the midgut environment.

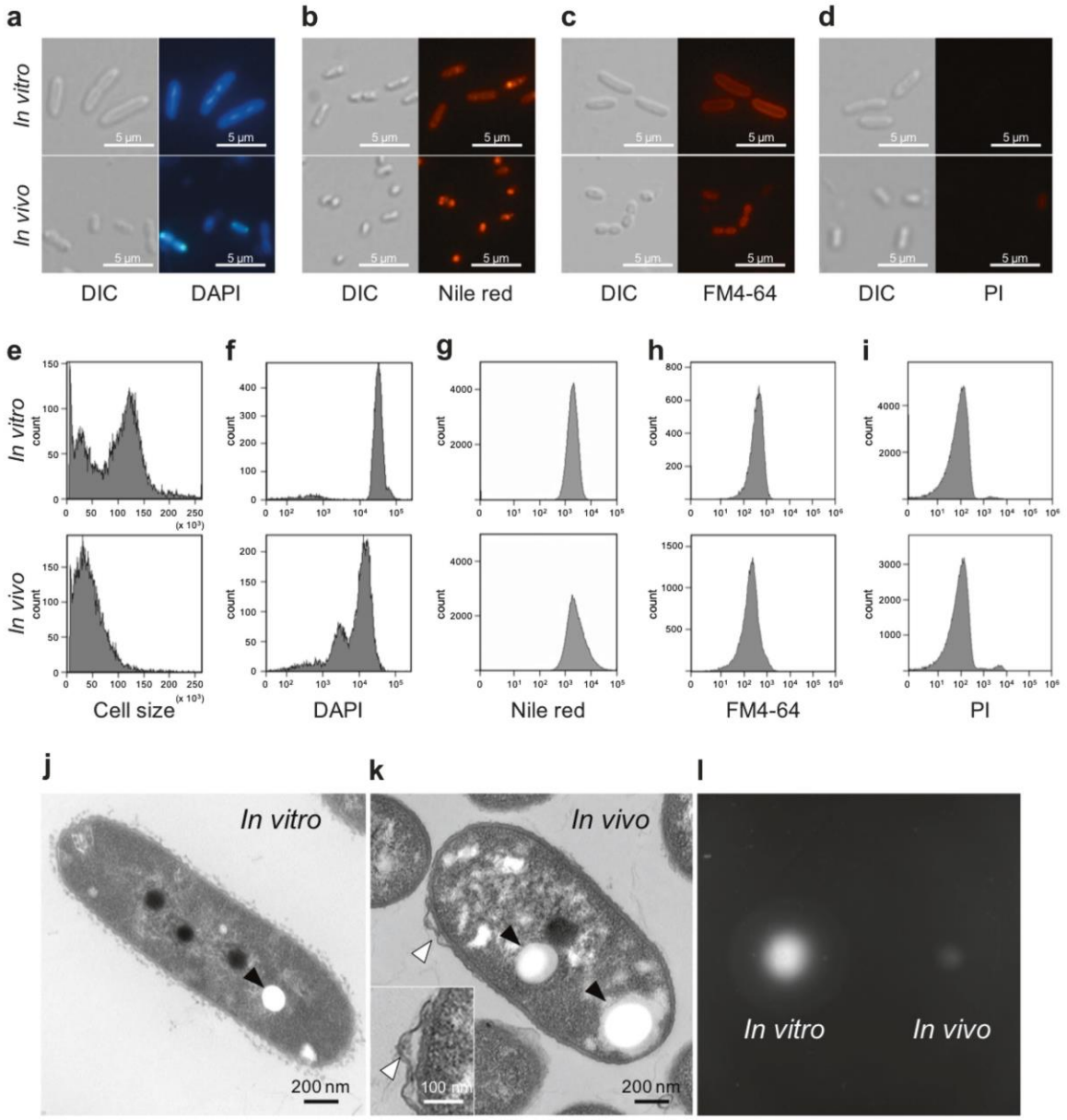


Fig. 1 Bacterial cell morphology of in vivo and in vitro *Burkholderia* symbiont cells. **a–d** Differential interference contrast (left) and fluorescence microscopy (right) images of in vitro and in vivo bacteria stained with DAPI **a**, Nile red **b**, FM4-64 **c**, and PI **d**. **e–i** Flow cytometry analysis of in vivo and in vitro symbiont cells measuring cell size by light scatter **e**, DNA content by DAPI staining **f**, PHA accumulation by Nile red staining **g**, and membrane area by FM4-64 staining **h**, membrane permeability by PI staining **i**. **j**, **k** Transmission electron microscopy images of an in vitro **j** and an in vivo **k** symbiont cell. Filled arrows indicate PHA granules and open arrows shows a membrane bleb. **l** Motility test of in vitro (left) and in vivo (right) cells (see also Supplementary Fig. S3)

in vivo *Burkholderia* symbionts are more susceptible to various environmental stresses

When symbiont cells were exposed to various antimicrobial compounds for a short time (1 h), in vivo symbionts were found to be more susceptible to membrane-disrupting agents, including SDS, proteinase K, ethanol, and cationic antimicrobial peptides (AMPs) (polymyxin B and LL-37), reactive oxygen (hydrogen peroxide), salt (NaCl), and high osmolarity (glucose) (Fig. 2; Supplementary Table S3). No obvious difference was detected between in vivo and in vitro cells in the

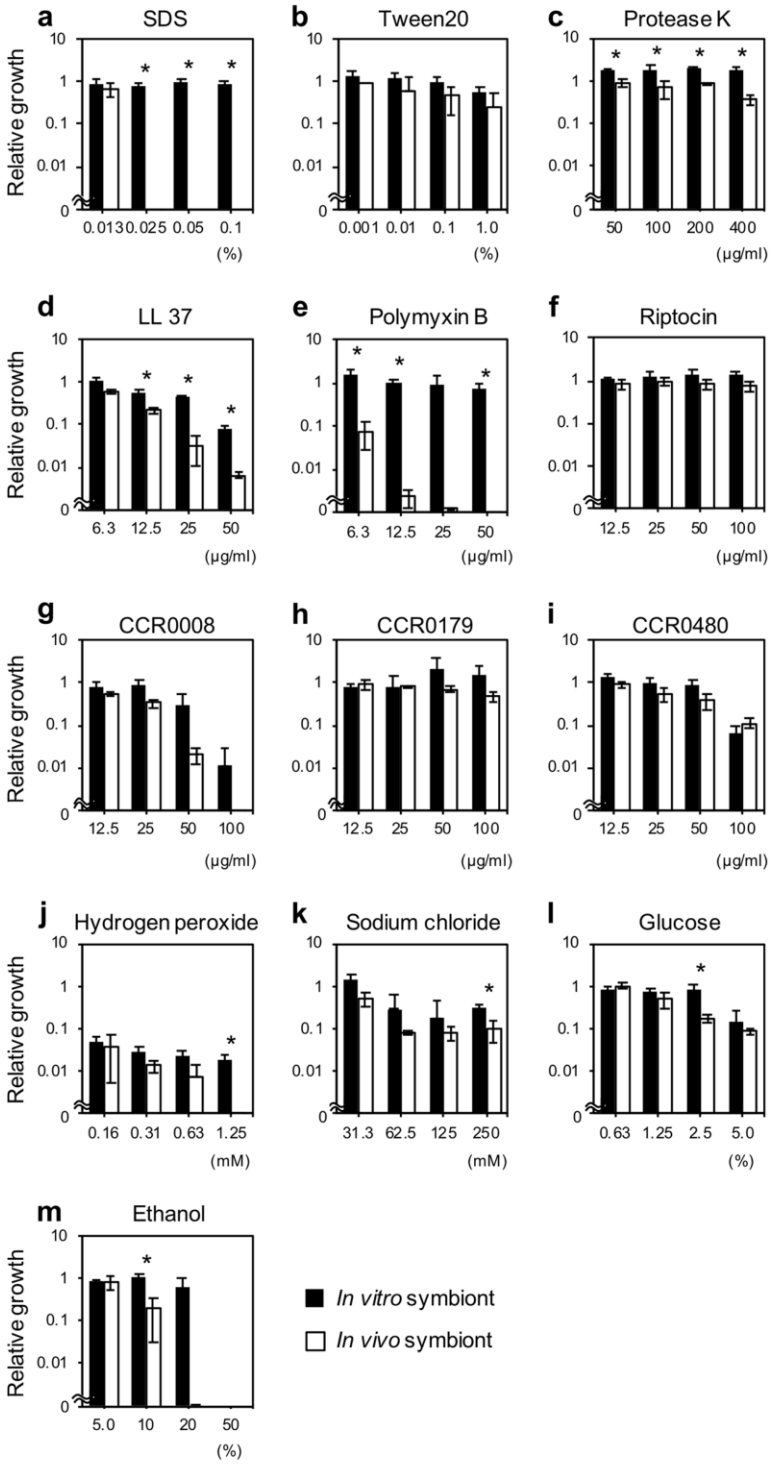


Fig. 2 Comparison of stress sensitivities between in vivo and in vitro bacteria. The in vivo and in vitro *Burkholderia* symbiont cells were exposed to **a** SDS, **b** tween 20, **c** protease K, **d** LL-37, **e** polymyxin B, **f** riptocin, **g** CCR0008, **h** CCR0179, **i** CCR0480, **j** hydrogen peroxide, **k** sodium chloride, **l** glucose, and **m** ethanol at the indicated concentrations. The relative growth (%) was normalized by setting untreated bacterial growth to 100%. Mean ± SD ($n = 3$) is shown in in vitro (solid bar) and in vivo cells (open bar). Statistically significant differences between in vivo and in vitro symbiont cells were analyzed by Student's *t*-test with Bonferroni correction: $^*P < 0.05$. CFU data of the stress tests are shown in Supplementary Table S3

presence of the non-ionic detergent Tween 20, and three cationic symbiotic AMPs (CCR0008, CCR0179, and CCR0480) (Fig. 2; Supplementary Table S3). The higher sensitivity to some of the AMPs and detergents, which are both stresses that target the cell membrane, strongly suggests cell membrane alterations in the in vivo symbionts, which is in agreement with the TEM observations (Fig. 1k; Supplementary Fig. S2).

RNA-sequencing of in vivo and in vitro *Burkholderia* symbionts

For a transcriptome comparison of the two *Burkholderia* cell-types, RNA-seq libraries were constructed from in vivo cells and 3 h-, 8 h-, and 16 h-cultured in vitro cells in triplicate (Supplementary Table S4). These in vitro cells corresponded to lag-phase, log-phase, and stationary-phase growing cells, respectively. More than 15,000,000 mapped reads were obtained from each in vitro sample, while over 5,800,000 mapped reads were obtained from each in vivo sample (Supplementary Table S4). In order to verify the RNA-seq data, qRT-PCR was performed on 8 randomly selected genes (Supplementary Table S5). The comparison of the qRT-PCR results with the RNA-seq data showed a high correlation coefficient value (R^2) > 0.9 for the fold change values of the in vivo expression versus the in vitro expression at 3 h, 8 h and 16 h of growth, respectively (Supplementary Fig. S4). These results demonstrate that the RNA-seq analysis provided overall reliable data.

General characteristics of the transcriptomics of in vivo symbiont cells

B. insecticola strain RPE64 has a multipartite genome composed of three chromosomes and two plasmids [45]. The overall expression level of genes located on these five genome components is not uniform (Fig. 3a; Supplementary Fig. S5). A principle component analysis (PCA) of the overall transcriptomes clustered the replicates of each condition demonstrating the high reproducibility of the transcriptomes (Fig. 3b). The PCA further revealed a strong partitioning of the conditions, where the transcriptome of the in vivo bacteria was unique compared to the three in vitro transcriptomes (Fig. 3b). The pairwise Pearson correlation coefficients revealed that the in vivo gene-expression profile is relatively similar to that of 3 h-cultured in vitro cells (Fig. 3c). The high correlation of the in vivo cells with lag- and log-phase cells and low correlation with stationary-phase cells is in large part driven by the expression of genes encoding core metabolisms such as cell division, DNA reprication, protein synthesis, and respiration (Fig. 3d).

Cell division, DNA replication, protein synthesis, and respiration

The genes involved in cell division and DNA replication were highly expressed in the in vivo cells as compared to in vitro stationary cells (Fig. 3d). Furthermore, the genes involved in protein synthesis (i.e. ribosomal subunits, translation factors, and protein chaperones such as GroEL/GroES and CspA mRNA chaperones) and respiration (i.e. NADH dehydrogenase, cytochrome c oxidase and ATP synthetase) were highly expressed in vivo to a similar level as exponentially growing in vitro cells (Fig. 3d). The overall data indicates the active proliferation of in vivo symbionts.

The expression pattern of all annotated genes encoding subunits of the respiratory chain complex I to IV and the ATP synthase and known markers of hypoxia (Supplementary Table S6) is most similar to 3 h cultures, suggesting that the respiratory activity of in vivo symbionts is high and that these bacteria are not limited by oxygen availability. The *cyoABCD* cluster is, however, very lowly expressed in the in vivo bacteria contrary to the culture conditions. Although it is commonly thought that Cyo is used at higher oxygen concentrations, there are also reports that indicate the upregulation and importance of Cyo in low oxygen concentrations [49], including in *Burkholderia pseudomallei* [50]. In addition, cytochrome bd is also very lowly expressed in in vivo bacteria and this terminal oxidase is known for its high affinity for O_2 and its activity under low oxygen concentrations [51]. Furthermore, a previous transcriptome study in *B. pseudomallei* [50] has identified several genes upregulated under hypoxia conditions, including genes for PHA synthesis, chemotaxis, motility, and some stress response proteins (*dps*, *hsp*, *usp*, and *clpAB*), none of which are induced in in vivo symbionts or many of which are even repressed compared to the culture condition (Supplementary Table S6).

Surface structure: LPS and cell-wall biosyntheses

In the LPS biosynthesis pathway, genes involved in lipid A, core oligosaccharide (core-OS), 4-amino-4-deoxy-L-arabinose (Ara4N) and O-antigen biosyntheses, and LPS transport were consistently expressed in in vivo cells (Supplementary Fig. S6a, Table S6). In the peptidoglycan biosynthesis pathway, most of the genes, including *murA-G*, *mraY*, *uppP*, *mviN*, *pbpB* and *mrdA*, were consistently expressed in vivo (Supplementary Fig. S6b, Table S6). The expression profiles of the LPS and peptidoglycan biosynthesis pathways strongly suggest that intact structures of LPS and cell wall are synthesized in the in vivo condition.

Identification of upregulated and downregulated genes in the in vivo symbiont

Next, we focused on specific differentially expressed genes between in vivo and in vitro symbiont cells. We compared the in vivo transcriptome with the 3 h-, 8 h-, and 16 h-cultured in vitro transcriptomes, and identified 527 in vivo-specific upregulated genes and 638 downregulated genes (Supplementary Fig. S7a-b, Tables S7, S8). The

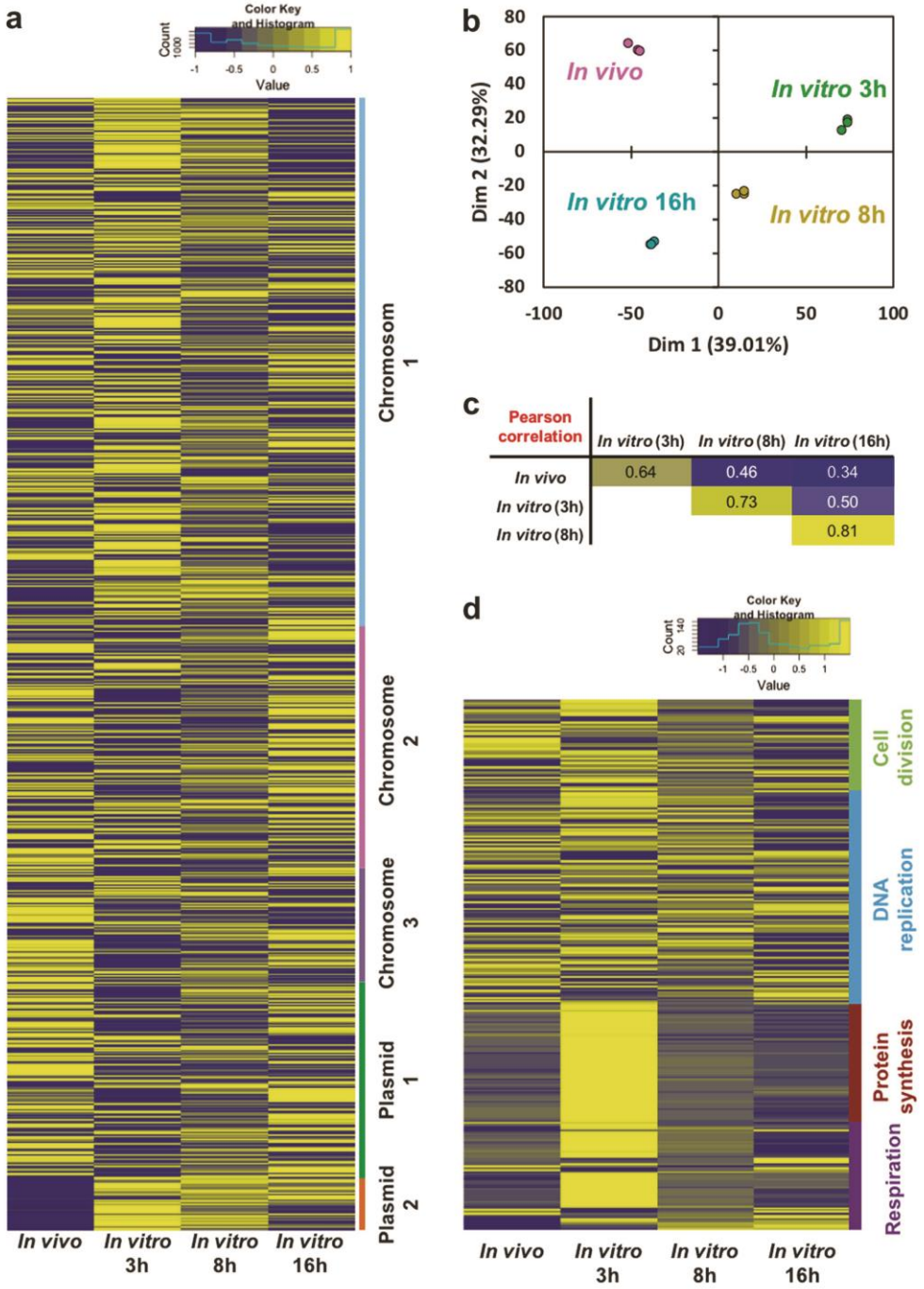


Fig. 3 Gene expression profiles of in vivo and in vitro *Burkholderia* symbiont. **a** heatmap of 5872 gene expression profiles, **b** Principal component analysis, and **c** the pairwise Pearson correlation coefficients of in vivo and 3 h-, 8 h-, and 16 h-cultured in vitro cells. In the heatmap, the normalized mean expression level of the three biological replicates was used, and the color scale from blue to yellow indicates the relative expression level. **d** Gene expression profile of core cellular functions. The heatmap shows the expressions of 281 genes encoding cell division, DNA replication, protein synthesis, and respiration in in vivo and in vitro symbiont cells

chromosome 2, chromosome 3, and plasmid 1 contained more in vivo upregulated genes, whereas chromosome 1 and plasmid 2 had more in vivo downregulated genes (Supplementary Fig. S7c). These in vivo upregulated and downregulated genes were classified into clusters of orthologous groups (COG) of the NCBI public database

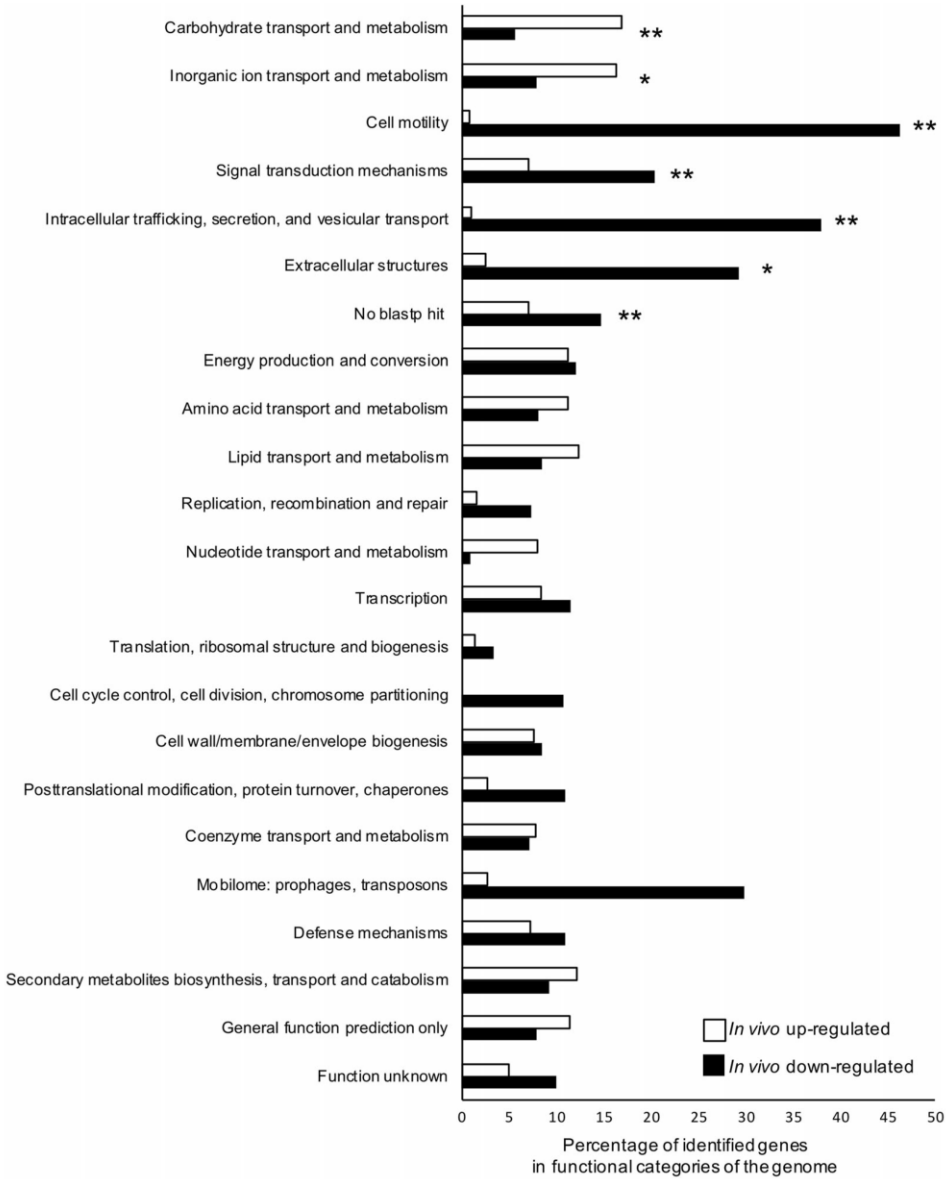


Fig. 4 COG functional classification of in vivo upregulated and down-regulated genes. The in vivo upregulated and down-regulated genes were classified using the NCBI COG 2014 database. Asterisks indicate statistically significant difference between in vivo and in vitro cells (Fisher's exact test with Bonferroni correction; $*P < 0.05$, $**P < 0.01$)

(Fig. 4). The in vivo upregulated genes were enriched mainly in two COG categories: "carbohydrate transport and metabolism" and "inorganic ion transport and metabolism". In contrast, the in vivo downregulated genes were significantly enriched in five COGs: "cell motility", "signal transduction mechanism", "intracellular trafficking", "secretion and vesicular transport", "extracellular structures", and "no blastp hit". In all COGs except cell motility, over 60% genes were equally expressed in in vivo and in vitro cells (Supplementary Fig. S8).

A detailed landscape of in vivo metabolic alterations was obtained by classifying in vivo upregulated and down-regulated genes with the KEGG (Kyoto Encyclopedia of Genes and Genomes) database. In the KEGG database, 204 in vivo upregulated genes and 196 down-regulated genes were annotated (Fig. 5; Supplementary Tables S7, S8). The in vivo upregulated and down-regulated genes/pathways are described in more detail in the following sections.

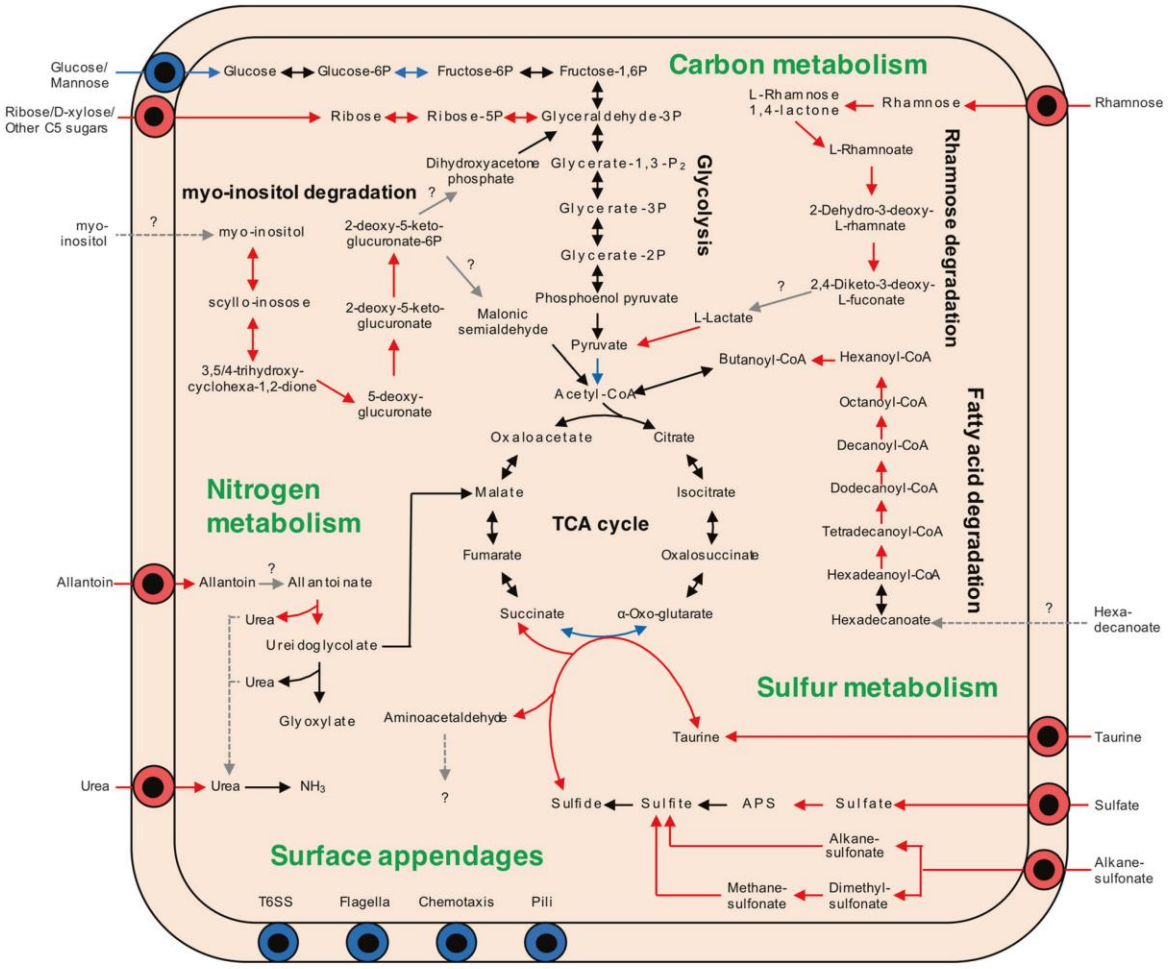


Fig. 5 Overview of KEGG metabolic pathways in in vivo upregulated and down-regulated genes. Arrows indicate metabolic direction, and the color shows gene expression levels in RNA-seq: red, in vivo upregulatd; blue, in vivo down-regulated; black, un-changed expression levels

Upregulated pathways in in vivo *Burkholderia* symbiont

Carbon metabolism

The transcriptome comparisons clearly demonstrated that in vivo cells strongly uptake and assimilate pentose sugars like ribose and rhammose, as well as myo-inositol as carbon sources (Fig. 5; Supplementary Table S6). On the other hand, the glucose transporters and the pathway upstream of glycolysis, which was highly expressed in vitro, were downregulated. Furthermore, in contrast to the down-regulation of the fatty acid biosynthesis pathway (Supplementary Table S6), the fatty acid degradation pathway was consistently and strongly upregulated in vivo (Fig. 5; Supplementary Table S6). These results strongly suggest that these sugars and fatty acid are carbon sources provided by the host insect.

Nitrogen metabolism

In the in vivo symbiont, transporters and assimilation pathways of allantoin and urea were specifically upregulated (Fig. 5; Supplementary Fig. S9a, Table S6). Most insects excrete uric acid as nitrogen waste and some species, including heteropteran bugs, convert it to allantoin or further to allantoic acid [52, 53]. In addition, although the ornitine cycle has not been determined in insects, urea is consistently detected in the hemolymph in diverse insects [54, 55]. The strong gene induction of the allantoin and urea pathways indicates that the *Burkholderia* symbiont utilizes these host nitrogen wastes inside the midgut crypts.

Sulfur metabolism

The in vivo symbiont highly and specifically expressed the sulfate reduction pathway and degradation pathway of sulfonates such as taurine and alkanesulphonate (Fig. 5; Supplementary Fig. S9b, Table S6). Inside the symbiont, sulfate could be assimilated to cysteine. Among the sulfur metabolism pathways, a taurine dioxigenase, *tauD*, was strongly upregulated (Supplementary Fig. S9b), which converts taurine and α-oxoglutarate to sulfite, succinate, and aminoacetaldehyde [56].

Other upregulated metabolisms are described in Supplementary Information.

Downregulated pathways in in vivo *Burkholderia* symbiont

The in vivo downregulated genes or pathways are involved in cell motility, chemotaxis, and type VI secretion system (see Supplementary Information). Remarkably, gene expression of the plasmid 2 was almost completely suppressed in vivo (Fig. 3a; Supplementary Fig. S5). We found that this was because of a frequent loss of this plasmid in the in vivo symbiont (see Supplementary Information).

Biosynthesis pathways of essential amino acids and B vitamins

We specifically analyzed biosynthesis genes of the essential amino acids and B vitamins, because these are potentially key metabolites in the stinkbug-*Burkholderia* symbiosis as it is the case in other nutrional symbioses of phytophagous insects. The biosynthesis pathways for the nine essential amino acids were not among the identified upregulated genes in the in vivo symbiont cells, but the genes were nevertheless highly expressed both in vivo and in vitro (Supplementary Fig. S10, Table S6). Similarly, biosynthesis pathways for B vitamins (i.e. thiamin, riboflavin, niacin, pantothenic acid, pyridoxine, cobalamin, folic acid, and biotin), including a few genes not annotated in the KEGG database, were highly expressed both in vivo and in vitro (Supplementary Fig. S11, Table S6).

Discussion

In general, the bacterial cell morphology strongly depends on environmental cues, such as the nutritional condition and AMP stress [57–59]. The smaller cell size and reduced DNA content of the in vivo symbiont (Fig. 1) could be interpreted as an indication that these bacteria are starved and that the midgut internal environment is nutritionally-poor or -biased for the symbiotic *Burkholderia* (Supplementary Fig. S1). However, the transcriptome analysis tells us a different message: the overall gene expression profile of the midgut bacteria is similar to the profile of exponentially growing in vitro cells rather than stationary phase ones, wherein genes involved in core cellular functions, such as cell division, protein biosyntheses, and respiration are highly expressed in the in vivo symbiont cells (Fig. 3d). The expression profiles further suggest that the symbiotic bacteria are not in a microaerobic environment. Thus, it is unlikely that nutrional or oxygen limitation provokes the specific morphology and cytology of the *Burkholderia* and therefore, other factors of the gut environment could determine the phenotype of the in vivo symbiont cells.

Morphological alterations have been described in other host-bacterium interactions. In the case of intracellular infection of macrophages by *Salmonella*, AMPs produced by the macrophages inhibit the division of the *Salmonella* cells and their continuous growth leads to filamentation of the bacteria [60]. Another well-known example is found in the legume-*Rhizobium* symbiosis, in which the *Rhizobium* cells become enlarged and polyploid because of the suppression of cell division by AMPs, called Nodule-specific Cysteine-Rich peptides (NCRs), inducing their transition into a unique cell-state called the bacteroid, which is specialized to fix atmospheric dinitrogen [61–63]. Interestingly, also the midgut crypts of *R. pedestris* produce AMP-like cationic peptides, called CCRs. More than 90 different genes encoding such CCR peptides are highly expressed by the midgut crypts and most of them are only induced when the crypts are colonized [39]. Although the biological function of the CCRs remains unclear, in this study we demonstrated that the chemically synthesized CCR0008 and CCR0480 have an antibiotic and growth inhibitory activity against the symbiont (Fig. 2; Supplementary Table S3), suggesting that CCRs may be involved in controlling the *Burkholderia* population in the gut. The production of this CCR peptide family by the midgut crypts is analogous to the NCR peptides produced in the legume-*Rhizobium* symbiosis and it will be of interest in the future to determine whether these peptides affect the bacterial cell cycle, DNA replication, or cell morphology.

Another notable feature of the in vivo symbiont is the high susceptibility to diverse stresses (Fig. 2). Previous studies have reported that cell surface structures are altered in in vivo *Burkholderia* symbionts and in particular that the O-antigen moiety of LPS is lost while the core-oligosaccharide structure is required to maintain a stable colonization [36, 37]. In addition, their cell envelope shows perturbations in the form of outer membrane blebs (Fig. 1k; Supplementary Fig. S2) [36]. These blebs could be emerging outer membrane vesicles (OMVs), which are commonly produced by Gram-negative bacteria enabling them

to deal with certain stress conditions [64]. Similar type of blebs and OMVs have been reported frequently to be produced by bacteria in response to antimicrobial peptides [65–69], including to the NCR peptides [62]. Thus, these blebs could be the result of the exposure of the in vivo symbiont to the CCR peptides. The susceptibility of the in vivo symbiont to diverse stresses could be largely explained by these surface alterations.

Although only the core-oligosaccharide is maintained in the in vivo symbiont and the O-antigen is absent [37], we found that both O-antigen and core-oligosaccharide biosynthesis pathway genes were expressed in the midgut bacteria (Supplementary Table S6). This contradiction between the transcriptome and biochemical data suggest that the synthesis of these surface structures is post-transcripionally affected by host factor(s). These could be glycolytic enzymes, inhibitors of O-antigen translation and transport, and/or AMPs attacking bacterial surface (such as CCRs) in the midgut crypts.

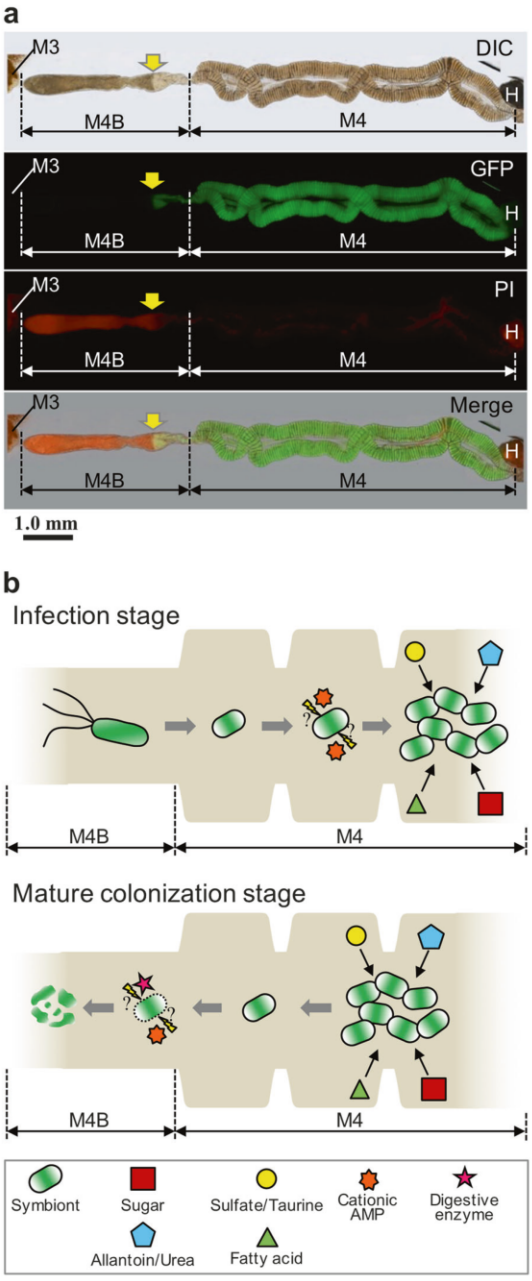
We further find that the gene cluster involved in the Ara4N modification of lipid A is constitutively expressed in all the tested conditions. This modification is known in some bacteria to be introduced onto lipid A upon sensing of AMPs [70–72], which renders lipid A less charged and therefore less vulnerable to AMPs. However, in *Burkholderia*, this modification is probably constitutively added to lipid A, which is essential for viability in *B. cenocepacia* [73]. If the same is true for *B. insecticola*, it is not surprising the constitutive expression of genes for the Ara4N modification of lipid A and therefore, it is unlikely that the increased sensitivity of the in vivo bacteria is due to the loss of the Ara4N modification.

The RNA-seq comparisons provided indications on the nutrients that are probided by the bean bug to the *Burkholderia* symbiont. In the midgut crypts, the host likely provides ribose, rhamnose, myo-inositol, and fatty acids as carbon sources, allantoin and urea as nitrogen sources, and sulfate, and sulfonates as sulfur sources (Fig. 5). Allantoin is the major nitrogen waste compound in heteropteran insects [52, 53], and urea is probably a nitrogen waste as well [54, 55]. Considering the highly upregulation of the allantoin and urea pathways (Fig. 5; Supplementary Fig. S9a), the recycling of nitrogen waste may be a biological function of the *Burkholderia* symbiont in the bean bug. Recycling of host metabolic waste is one of the common traits in insect symbionts [6, 74]. Such a "recycling" symbiosis could enable the host to more efficiently utilize nutrients from food and to save energy, resulting in the more rapid growth and larger body size of the symbiotic bean bugs.

The other intriguing observation is the sulfer metabolism of the *Burkholderia* symbiont. In general, insects do not have a sulfate-reducing enzyme and sulfate has to be reduced to sulfite that can be assimilated [75], while our RNA-seq results revealed that the in vivo symbiont can assimilate sulfate to cysteine. It is notable that genes involved in sulfate assimilation have been reported in other symbiotic bacteria in insects [76–79], suggesting the sulfate assimilation is a common feature of insect symbionts. Furthermore, another notable point here is that the taurine degradation is tightly linked to the TCA cycle, wherein *tauD* provides succinate by oxidising taurine and α-oxoglutarate [56]. Since the genes involved in the conversion from α-oxoglutarate to succinate in the TCA cycle are strongly suppressed in vivo (Fig. 5; Supplementaly Table S6), the taurine metabolism appears likely to complement this conversion process, which may enable the host to control the symbiont via taurine feeding.

In many insects, the nutritional benefit of symbionts is the provisioning of essential amino acids and/or B vitamins [80, 81]. A recent physiological study on the gut bacteria of a cotton pest, *Dysdercus fasciatus* (Heteroptera: Pyrrhocoridae), has revealed that the symbiotic bacteria provide B vitamins [82], which are deficient in cotton seeds. Unlike cotton seeds, soybean seeds, the food of *R. pedestris*, are known to be rich in B vitamins except vitamin B_{12} (cobalamin) (Supplementary Fig. S12) [83, 84]. Biosynthetic pathways for all 8 kinds of B vitamins were highly expressed in the in vivo condition, suggesting the possibility that the symbiont provide the host insect with B vitamins, particularly vitamin B_{12}. Although soybean seeds are rich in amino acids, two essential amino acids, methionine, and tryptophan, are limited (Supplementary Fig. S12) [83, 84]. The *Burkholderia* symbiont has retained intact pathways for producing these essential amino acids and the involved genes were highly expressed in vivo, suggesting that the *Burkholderia* symbiont mainly supplies methionine and tryptophan to the host.

How exactly does the host extract the nutrients derived from the *Burkholderia* symbiont? It seems that an excess of symbionts in the midgut crypts is not excreted since *R. pedestris* excrements do not contain the symbiont [25], but that they instead flow into the M4B region [40], wherein AMPs and digestive enzymes like cathepsins are highly and specifically expressed [39] and kill the bacteria [40, 41]. Furthermore, in dissected midguts infected with GFP-expressing *Burkholderia* and stained with PI, we noticed that the bacteria in the M4B region close to the M4 are still alive, GFP-positive and PI-negative, while in the more distant regions of M4B the GFP-signal disappears suddenly and PI staining becomes strong (Fig. 6a). This suggests that whole symbiont cells are digested in the M4B region and that the liberated nutrients are absorbed by the host insect. Based on these biological features of the in vivo symbiont cells, we here propose the following in vivo life cycle of the *Burkholderia* symbiont (Fig. 6b): (1) after colonization in



◀**Fig. 6** A hypothetical model for the life cycle of the *Burkholderia* symbiont in the midgut of *R. pedestris*. **a** The PI staining in the midgut of 5th instar nymph infected with the strain RPE225 (a GFP-expressing mutant). DIC, GFP fluorescence, PI fluorescence, and merged images are shown. Arrows indicate the border of GFP and PI fluorescent signals. M3, midgut 3rd section; M4, midgut 4th section (crypts); M4B, M4 bulb; H, hindgut. A red auto-fluorescence was observed in M3 and H. **b** Graphical summary of *Burkholderia* features and midgut functions of *R. pedestris*. (Upper) During the infection stage of *Burkholderia* symbiont at the insect midgut, the *Burkholderia* symbiont cells lose flagellar motility and modify their envelope under the influence of host stress factors such as cationic AMPs (e.g. CCRs), they proliferate by metabolizing host waste materials, such as sulfate and allantoin. (Lower) In the mature colonization stage, these symbiont cells are digested in the M4B section mediated by host factors such as cathepsin proteases and CCRs, and the host absorbs nutrients derived from whole-bacterial cells

the midgut crypts, the symbiont actively proliferates by use of host-provided nutrients including metabolic wastes such as allantoin; (2) excess symbiont cells flow into M4B and are digested by the host, wherein the backward flow of the symbiotic bacteria from M4 to M4B could be made possible by a temporal and local regain of motility of the symbiont, by pressure generated by the multiplication of the symbiont in M4, or by gut peristaltic movements; (3) not a specific nutrient but the ensemble of nutrients of the symbiont cells is absorbed, contributing to the rapid growth and enhanced reproduction of the bean bug host. Such a symbiotic association, i.e. a host cultivates and consumes whole cells of its symbiont, has been reported in other animal-microbe symbiotic systems, such as the cultivation symbiosis in leaf-cutting ants [85] and the gut symbiosis in herbivorous rodents that show a coprophagy behavior [86, 87].

Conclusion and perspective

In this study, we demonstrated how the *Burkholderia* symbiont copes with the environment of the midgut crypts, and suggested what roles the symbiont plays in the symbiotic organ. To understand the host-symbiont nutritional interactions in the stinkbug-*Burkholderia* association more deeply, metabolomics of the midgut crypts and feces, and/or isotope-labeling and metabolite-tracing experiments will be needed. These experiments would also elucidate the mechanisms that enable the bean bug to obtain a bigger body size and produce a higher egg number in the presence of the *Burkholderia* symbionts. Furthermore, it should be noted that almost 25% of the in vivo upregulated genes are functionally unknown, which underscores how much the endosymbiosis between the bean bug and *Burkholderia* is still enigmatic. Nevertheless, this report about the bacterial features in the bean bug is important to orient future work on this model symbiosis.

Acknowledgements Part of this work was performed at the Open Facility of Hokkaido University and the Imagerie-Gif facilities. This work has benefited from the expertise of the High-throughput Sequencing Platform of I2BC and from support by the Labex 'Saclay Plant Sciences' (ANR-11-IDEX-0003-02). This work was supported by the MEXT KAKENHI (15H05638 to Y.K.), the JSPS Research Fellowship for Young Scientists (14J03996 and 20170267 to T.O.) and the JSPS Invitation Fellowship for Research in Japan (L-14556 to P.M.).

Author contributions TO, PM, and YK designed the project. TO, TS, and PM performed the stress tests. MT and QB conducted flow cytometry analyses. XYM and YM performed electron microscopy. TO, RF, SS, TF, FL, KT, and PM conducted RNA-seq and analyzed transcriptome data. TO, PM, and YK wrote the manuscript. All co-authors edited the manuscript before submission.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Schwartzman JA, Ruby EG. Stress as a normal cue in the symbiotic environment. Trends Microbiol. 2016;24:414–24.
2. Yang DC, Blair KM, Salama NR. Staying in shape: the impact of cell shape on bacterial survival in diverse environments. Microbiol Mol Biol Rev. 2016;80:187–203.
3. Buchner P. Endosymbiosis of animals with plant microorganims. New York: Interscience; 1965.
4. Bourtzis K, Miller T. Insect symbiosis. Boca Raton, FL: CRC Press; 2003.
5. Kikuchi Y. Endosymbiotic bacteria in insects: their diversity and culturability. Microbes Environ. 2009;24:195–204.
6. Ankrah NYD, Douglas AE. Nutrient factories: metabolic function of beneficial microorganisms associated with insects. Environ Microbiol. 2018;20:2002–11.
7. McCutcheon JP, Moran NA. Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol. 2011;10:13–26.
8. Utami YD, Kuwahara H, Murakami T, Morikawa T, Sugaya K, Kihara K, et al. Phylogenetic diversity and single-cell genome analysis of "*Melainabacteria*", a non-photosynthetic cyanobacterial group, in the termite gut. Microbes Environ. 2018;33:50–57.
9. Noda S, Shimizu D, Yuki M, Kitade O, Ohkuma M. Host-symbiont cospeciation of termite-gut cellulolytic protists of the genera *Teranympha* and *Eucomonympha* and their *Treponema* endosymbionts. Microbes Environ. 2018;33:26–33.
10. Salem H, Bauer E, Kirsch, Berasategui A, Cripps M, Weiss B, et al. Drastic genome reduction in an herbivore's pectinolytic symbiont. Cell. 2017;171:1520–31.
11. Sabree ZL, Kambhampati S, Moran NA. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. Proc Natl Acad Sci USA. 2009;106:19521–6.
12. Kaltenpoth M. Actinobacteria as mutualists: general healthcare for insects? Trends Microbiol. 2009;17:529–35.
13. Flórez LV, Biedermann PH, Engl T, Kaltenpoth M. Defensive symbioses of animals with prokaryotic and eukaryotic microorganisms. Nat Prod Rep. 2015;32:904–36.
14. Werren JH, Baldo L, Clark ME. *Wolbachia*: master manipulators of invertebrate biology. Nat Rev Microbiol. 2008;6:741.
15. Itoh H, Tago K, Hayatsu M, Kikuchi Y. Detoxifying symbiosis: microbe-mediated detoxification of phytotoxins and pesticides in insects. Nat Prod Rep. 2018;35:434–54.
16. Bright M, Bulgheresi S. A complex journey: transmission of microbial symbionts. Nat Rev Microbiol. 2010;8:218–30.
17. Baumann P. Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. Annu Rev Microbiol. 2005;59:155–89.
18. Mergaert P, Kikuchi Y, Shigenobu S, Nowack ECM. Metabolic integration of bacterial endosymbionts through antimicrobial peptides. Trends Microbiol. 2017;25:703–12.
19. Kadosawa T, Santa H. Growth and reproduction of soybean pod bugs (Heteroptera) on seeds of legumes. Bull Chugoku Nat Agric Exp Stn. 1981;19:75–97.
20. Kono S. Ecological studies of stink bugs injuring soybean seeds for developing effective control measures. Sp Rep Hyogo Agric Exp Stn. 1991;16:32–68.
21. Schaefer CW, Panizzi AR. Heteroptera of economic importance. Boca Raton, FL: CRC press; 2000.
22. Kikuhara Y. The Japanese species of the genus *Riptortus* (Heteroptera, Alydidae) with description of a new species. Jpn J Syst Entomol. 2005;11:299–311.
23. Takeshita K, Tamaki H, Ohbayashi T, Meng XY, Sone T, Mitani Y, et al. *Burkholderia insecticola* sp. nov., a gut symbiotic bacterium of the bean bug *Riptortus pedestris*. Int J Syst Evol Microbiol. 2018;68:2370–4.
24. Takeshita K, Kikuchi Y. *Riptortus pedestris* and *Burkholderia* symbiont: an ideal model system for insect-microbe symbiotic associations. Res Microbiol. 2016;168:175–87.
25. Kikuchi Y, Hosokawa T, Fukatsu T. Insect-microbe mutualism without vertical transmission: a stinkbug acquires a beneficial gut symbiont from the environment every generation. Appl Environ Microbiol. 2007;73:4308–16.
26. Kikuchi Y, Fukatsu T. Live imaging of symbiosis: spatiotemporal infection dynamics of a GFP-labelled B*urkholderia* symbiont in the bean bug *Riptortus pedestris*. Mol Ecol. 2014;23:1445–56.
27. Lee JB, Park KE, Lee SA, Jang SH, Eo HJ, Jang HA, et al. Gut symbiotic bacteria stimulate insect growth and egg production by modulating hexamerin and vitellogenin gene expression. Dev Comp Immunol. 2017;69:12–22.
28. Kikuchi Y, Hosokawa T, Fukatsu T. Specific developmental window for establishment of an insect-microbe gut symbiosis. Appl Environ Microbiol. 2011;77:4075–81.
29. Kuechler SM, Matsuura Y, Dettner K, Kikuchi Y. Phylogenetically diverse *Burkholderia* associated with midgut crypts of spurge bugs, *Dicranocephalus* spp. (Heteroptera: Stenocephalidae). Microbes Environ. 2016;31:145–53.
30. Kikuchi Y, Yumoto I. Efficient colonization of the bean bug *Riptortus pedestris* by an environmentally transmitted *Burkholderia* symbiont. Appl Environ Microbiol. 2013;79:2088–91.
31. Kim JK, Lee BL. Symbiotic factors in *Burkholderia* essential for establishing an association with the bean bug, *Riptortus pedestris*. Arch Insect Biochem Physiol. 2015;88:4–17.
32. Ohbayashi T, Takeshita K, Kitagawa W, Nikoh N, Koga R, Meng XY, et al. Insect's intestinal organ for symbiont sorting. Proc Natl Acad Sci USA. 2015;112:E5179–E5188.
33. Kinosita Y, Kikuchi Y, Mikami N, Nakane D, Nishizaka T. Unforeseen swimming and gliding mode of an insect gut symbiont, *Burkholderia* sp. RPE64, with wrapping of the flagella around its cell body. ISME J. 2018;12:838–48.
34. Kim JK, Lee HJ, Kikuchi Y, Kitagawa W, Nikoh N, Fukatsu T, et al. Bacterial cell wall synthesis gene *uppP* is required for *Burkholderia* colonization of the stinkbug gut. Appl Environ Microbiol. 2013;79:4879–86.
35. Kim JK, Won YJ, Nikoh N, Nakayama H, Han SH, Kikuchi Y, et al. Polyester synthesis genes associated with stress resistance are involved in an insect-bacterium symbiosis. Proc Natl Acad Sci USA. 2013;110:E2381–2389.
36. Kim JK, Son DW, Kim CH, Cho JH, Marchetti R, Silipo A, et al. Insect gut symbiont susceptibility to host antimicrobial peptides caused by alteration of the bacterial cell envelope. J Biol Chem. 2015;290:21042–53.
37. Kim JK, Jang HA, Kim MS, Cho JH, Lee JB, Lorenzo FD, et al. The lipopolysaccharide core oligosaccharide of *Burkholderia* plays a critical role in maintaining a proper gut symbiosis with the bean bug *Riptortus pedestris*. J Biol Chem. 2017;292:19226–37.

38. Jang SH, Jang HA, Lee J, Kim JU, Lee SA, Park KE, et al. PhaR, a negative regulator of PhaP, modulates the colonization of a *Burkholderia* gut symbiont in the midgut of the host insect, *Riptortus pedestris*. Appl Environ Microbiol. 2017;83:e00459–17.
39. Futahashi R, Tanaka K, Tanahashi M, Nikoh N, Kikuchi Y, Lee BL, et al. Gene expression in gut symbiotic organ of stinkbug affected by extracellular bacterial symbiont. PLoS ONE. 2013;8: e64557.
40. Kim JK, Kim NH, Jang HA, Kikuchi Y, Kim CH, Fukatsu T, et al. Specific midgut region controlling the symbiont population in an insect-microbe gut symbiotic association. Appl Environ Microbiol. 2013;79:7229–33.
41. Byeon JH, Seo ES, Lee JB, Lee MJ, Kim JK, Yoo JW, et al. A specific cathepsin-L-like protease purified from an insect midgut shows antibacterial activity against gut symbiotic bacteria. Dev Comp Immunol. 2015;53:79–84.
42. Jang HA, Seo ES, Seong MY, Lee BL. A midgut lysate of the *Riptortus pedestris* has antibacterial activity against LPS O-antigen-deficient *Burkholderia* mutants. Dev Comp Immunol. 2016;67:97–106.
43. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17:10–12.
44. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009;25:1754–60.
45. Shibata TF, Maeda T, Nikoh N, Yamaguchi K, Oshima K, Hattori M, et al. Complete genome sequence of *Burkholderia s*p. strain RPE64, bacterial symbiont of the bean bug *Riptortus pedestris*. Genome Announc. 2013;1:e00441–13.
46. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
47. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Softw. 2008;25:1–18.
48. Lee JB, Byeon JH, Jang HA, Kim JK, Yoo JW, Kikuchi Y, et al. Bacterial cell motility of *Burkholderia* gut symbiont is required to colonize the insect gut. FEBS Lett. 2015;589:2784–90.
49. Lunak ZR, Noel KD. A quionol oxidase, encoded by *cyoABCD*, is utilized to adapt to lower O_2 concentrations in *Rhizobium etli* CFN42. Microbiol . 2015;161:203–12.
50. Hamad MA, Austin CR, Stewart AL, Higgins M, Vazquez-Torres A, Voskuil MI. Adaptation and antibiotic tolerance of anaerobic *Burkholderia pseudomallei*. Antimicrob Agents Chemother. 2011;55:3313–23.
51. Borisov VB, Gennis RB, Hemp J, Verkhovsky MI. The cytochrome *bd* respiratory oxygen reductases. Biochim Biophys Acta. 2011;1807:1398–413.
52. Bursell E. The excretion of nitrogen in insects. Adv Insect Physiol. 1967;4:33–67.
53. Cochran DG. Nitrogen excretion in cockroaches. Annu Rev Entomol. 1985;30:29–49.
54. Corrigan JJ. Nitrogen metabolism in insects. In: Campbell JW editors. *The Comparative Biochemistry of Nitrogen Metabolism: The invertebrates*. 1st edn. New York, USA: Academic Press; 1970. p. 388–488.
55. Mullins D. Chemistry and physiology of the hemolymph. In: Kerkut GA, Gilbert LI editors. *Comprehensive Insect Physiology, Biochemistry and Pharmacology*. 3rd edn. London, UK: Pergamon; 1985. p. 356–400.
56. Eichhorn E, van der Ploeg JR, Kertesz MA, Leisinger T. Characterization of α-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. J Biol Chem. 1997;272:23031–6.
57. Young KD. The selective value of bacterial shape. Microbiol Mol Biol Rev. 2006;70:660–703.
58. Young KD. Bacterial morphology: why have different shapes? Curr Opin Microbiol. 2007;10:596–600.
59. Willis L, Huang KC. Sizing up the bacterial cell cycle. Nat Rev Microbiol. 2017;15:606.
60. Rosenberger CM, Gallo RL, Finlay BB. Interplay between antibacterial effectors: a macrophage antimicrobial peptide impairs intracellular *Salmonella* replication. Proc Natl Acad Sci USA. 2004;101:2422–7.
61. Czernic P, Gully D, Cartieaux F, Moulin L, Guefrachi I, Patrel D, et al. Convergent evolution of endosymbiont differentiation in dalbergioid and inverted repeat-lacking clade legumes mediated by nodule-specific cysteine-rich peptides. Plant Physiol. 2015;169:1254–65.
62. Montiel J, Downie JA, Farkas A, Bihari P, Herczeg R, Balint B, et al. Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. Proc Natl Acad Sci USA. 2017;114:5041–6.
63. Mergaert P. Role of antimicrobial peptides in controlling symbiotic bacterial populations. Nat Prod Rep. 2018;35:336–56.
64. Schwechheimer C, Kuehn MJ. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. Nat Rev Microbiol. 2015;13:605–19.
65. Manning AJ, Kuehn MJ. Contribution of bacterial outer membrane vesicles to innate bacterial defense. BMC Microbiol. 2011;11:258.
66. Duperthuy M, Sjöström AE, Sabharwal D, Damghani F, Uhlin BE, Wai SN. Role of the *Vibrio cholerae* matrix protein Bap1 in cross-resistance to antimicrobial peptides. PLoS Pathog. 2013;9: e1003620.
67. Kulkarni HM, Swamy ChVB, Jagannadham MV. Molecular characterization and functional analysis of outer membrane vesicles from the Antarctic bacterium *Pseudomonas syringae* suggest a possible response to environmental conditions. J Proteome Res. 2014;13:1345–1358.
68. Chileveru HR, Lim SA, Chairatana P, Wommack AJ, Chiang I-L, Nolan EM. Visualizing attack of *Escherichia coli* by the antimicrobial peptide Human Defensin 5. Biochemistry. 2015;54:1767–1777.
69. Urashima A, Sanou A, Yen H, Tobe T. Enterohaemorrhagic *Escherichia coli* produces outer membrane vesicles as an active defence system against antimicrobial peptide LL-37. Cell Microbiol. 2017;19:e12758.
70. Raetz CRH, Reynolds CM, Trent MS, Bishop RE. Lipid A modification systems in Gram-negative bacteria. Annu Rev Biochem. 2007;76:295–329.
71. Chen HD, Groisman EA. The biology of the PmrA/PmrB two-component system: the major regulator of lipopolysaccharide modifications. Annu Rev Microbiol. 2013;67:83–112.
72. Dalebroux ZD, Miller SI. Salmonellae PhoPQ regulation of the outer membrane to resist innate immunity. Curr Opin Microbiol. 2014;17:106–13.
73. Ortega XP, Cardona ST, Brown AR, Loutet SA, Flannagan RS, Campopiano DJ, et al. A putative gene cluster for aminoarabinose biosynthesis is essential for *Burkholderia cenocepacia* viability. J Bacteriol. 2007;189:3639–44.
74. Kashima T, Nakamura T, Tojo S. Uric acid recycling in the shield bug, *Parastrachia japonensis* (Hemiptera: Parastrachiidae), during diapause. J Insect Physiol. 2006;52:816–25.
75. Zientz E, Dandekar T, Gross R. Metabolic interdependence of obligate intracellular bacteria and their insect hosts. Microbiol Mol Biol Rev. 2004;68:745–70.
76. Patino-Navarrete R, Moya A, Latorre A, Pereto J. Comparative genomics of *Blattabacterium cuenoti*: the frozen legacy of an ancient endosymbiont genome. Genome Biol Evol. 2013;5:351–61.
77. Gil R, Silva FJ, Zientz E, Delmotte F, González-Candelas F, Latorre A, et al. The genome sequence of *Blochmannia*

floridanus: comparative analysis of reduced genomes. Proc Natl Acad Sci USA. 2003;100:9388–93.
78. Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature. 2000;407:81–86.
79. Santos-Garcia D, Latorre A, Moya A, Gibbs G, Hartung V, Dettner K, et al. Small but powerful, the primary endosymbiont of moss bugs, Candidatus *Evansia muelleri*, holds a reduced genome with large biosynthetic capabilities. Genome Biol Evol. 2014;6:1875–93.
80. Douglas AE. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. Annu Rev Entomol. 1998;43:17–37.
81. Engel P, Moran NA. The gut microbiota of insects-diversity in structure and function. FEMS Microbiol Rev. 2013;37:699–735.
82. Salem H, Bauer E, Strauss AS, Vogel H, Marz M, Kaltenpoth M. Vitamin supplementation by gut symbionts ensures metabolic homeostasis in an insect host. Proc R Soc Lond B Biol Sci. 2014;281:20141838.
83. Sharma S, Kaur M, Goyal R, Gill BS. Physical characteristics and nutritional composition of some new soybean *(Glycine max* (L.) Merrill) genotypes. J Food Sci Technol. 2014;51:551–7.
84. Edelman M, Colt M. Nutrient value of leaf vs. seed. Front Chem. 2016;4:32.
85. Mehdiabadi NJ, Schultz TR. Natural history and phylogeny of the fungus-farming ants (Hymenoptera: Formicidae: Myrmicinae: Attini). Myrmecol News. 2010;13:37–55.
86. Torrallardona D, Harris CI, Fuller MF. Microbial amino acid synthesis and utilization in rats: the role of coprophagy. Br J Nutr. 1996;76:701–9.
87. Sakaguchi E. Digestive strategies of small hindgut fermenters. Anim Sci J. 2003;74:327–37.

Supplementary Information of:

Comparative cytology, physiology and transcriptomics of free-living and symbiotic *Burkholderia* reveal features for colonization of its insect host

by Ohbayashi *et al.*

Supplementary Materials and Methods

Fluorescence microscopy

Bacterial cells were stained with 50 µg/ml of 4,6-diamidino-2-phenylindole (DAPI) or propidium iodide (PI). Cells were also stained with 5 µg/ml of N-(3-Triethyl-ammoniumpropyl)-4-(6-(4-(diethylamino) phenyl) hexatrienyl) pyridinium (FM4-64) and 0.1 % nile red for visualizing cell membranes and PHA accumulation, respectively. In DAPI and FM4-64 staining, the symbiont cells were incubated at 60 °C for 10 min prior to staining, and live cells were used in PI and nile red staining. DAPI, PI, and FM4-64 staining were performed for 1 minute at room temperature. After nile red staining for 15 minutes at room temperature, the cells were centrifuged at 15,000 rpm for 1 min, washed twice with PBS buffer and 8% acetic acid was added. The cells were further incubated for 1 min, centrifuged at 15,000 rpm for 1 min, and then washed three times with PBS buffer. The cells stained with the different dyes were investigated with a fluorescence microscope, DMI 4000B (Leica Ltd, Cambridge, UK). Observations were made on three biological replicates.

Flow cytometry

Bacterial cells stained with fluorescent dyes as above were diluted to suspensions with an optical density (OD_{600}) of 0.01 to 0.04, and analyzed on a BD Aria II (BD Bioscience, San Jose, CA) flow cytometer with excitation at 375 nm for DAPI staining, or on a Cytoflex (Beckman Coulter, Fullerton, FL, USA) cytometer with excitation at 638 nm for PI, nile red, and FM4-64 staining. All flow cytometry measurements were performed

independently at least three times, each with similar results.

Electron microscopy

Cells were prefixed with 2.5% glutaraldehyde in 0.1 M sodium phosphate buffer (pH 7.4) at 4°C overnight, and then post-fixed in 2% osmium tetroxide at 4°C for 60 min. After a series of dehydration steps with ethanol, the samples were embedded in Epon 812 resin (TAAB Ltd, Berkshire, UK). Ultrathin sections were cut by an ultramicrotome (EM UC7, Leica Ltd, Cambridge, UK), mounted on a copper mesh, stained with uranyl acetate and lead citrate, and then observed under a transmission electron microscope (H-7600, Hitachi Ltd, Tokyo, Japan).

Motility test

The *in vivo* and *in vitro* cells were adjusted to an OD_{600} of 1.0 in 1.0 ml of PBS buffer. One µl of each bacterial suspension was inoculated onto semisolid YG agar [0.4% (wt/vol) agar in YG medium], incubated at 27°C for 16 to 24 hours. Bacterial motility was estimated by determination of the diameter of bacterial halo with six biological replicates.

Quantitative reverse-transcription PCR (qRT-PCR)

Independent total RNA samples were extracted using the procedure as above for the RNA-seq experiments. The cDNA was synthesized with the RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific, Wilmington, NC, USA) using 500 ng of

total RNA as a template and then subjected to real-time quantitative PCR using the FastStart Essential DNA Green Master (Roche Scientific, Indianapolis, IN, USA) and 0.50 μM of primers. Primer sequences are provided in Supplementary table S4. The qPCR reactions were run on a LightCycler® 96 System (Roche Scientific, Indianapolis, IN, USA). The PCR temperature profile was set to 95 °C for 10 minutes, 40 cycles of 95 °C for 15 seconds, 55 °C for 15 seconds and 72 °C for 15 seconds, and then 95 °C for 1 seconds and 40 °C for 20 seconds. The comparative C_T ($\Delta\Delta C_T$) method was used to calculate relative gene expression levels with *opcP*, a house-keeping gene of *Burkholderia* sp. RPE64, as a reference. Observations were made on three biological replicates.

Statistical analyses

The statistical analyses in this study were performed by use of the R software (ver. 2.15.2, http://www.R-project.org). Results of the stress sensitivity tests, RNA-seq and COG functional classifications were analyzed by Student's *t*-test with a Bonferroni correction, FDR analysis and Fisher's exact test with Bonferroni correction, respectively.

Supplementary Results

The cell size and DNA content of *in vivo* and *in vitro* bacteria shift with growth phase and culture media

Flow cytometry analysis of DAPI stained bacteria showed that the *in vivo* cells had a

lower DNA content than the *in vitro* grown symbiont cells (Supplementary Fig. S1). It is known that in different bacteria, the nutritional quality of the environment has a profound impact on the cell size and that these bacteria grown under nutritional limitation are smaller than when grown in rich media (Willis and Huang, 2017). To test if this is also the case in the *Burkholderia* symbiont, we analyzed the cell size and DNA content in exponentially growing and stationary phase cultures in rich media and in minimal media. In the rich YG medium, bacteria were large and had an 8C DNA content when growing in the exponential phase but they gradually shifted to smaller size and lower DNA content when reaching the stationary phase (Supplementary Fig. S1). Also when the bacteria where grown in the minimal medium MM in which their doubling time is 3 hours instead of 2 hours in YG medium, bacteria where smaller and had lower DNA content which further shifted to lower values in stationary phase (Supplementary Fig. S1). Thus also in this *Burkholderia* strain, the size of the cells and their DNA content correlate with the nutritional status of the environment. Therefore, a possible explanation for the small cells with low DNA content in the *in vivo Burkholderia* is a nutritional limitation.

Expression of specific pathways in *in vivo Burkholderia*

[*Oxidative phosphorylation*]

Among a total of 60 genes of the oxidative phosphorylation pathway, no up-regulated and 13 down-regulated genes were identified in the *in vivo* symbiont (Supplementary table S7). Particularly, genes encoding NADH dehydrogenase, cytochrome c oxidase

and ATP synthetase were highly expressed in 3h-cultured *in vitro* cells.

[*Cell motility and chemotaxis*]

Flagellar motility and chemotaxis pathways were drastically down-regulated in *in vivo* symbionts. Among 41 genes for flagellar motility, only no genes were up-regulated, whereas 35 other genes were down-regulated in *in vivo* cells (Supplementary table S7). Similarly, among 45 genes encoding bacterial chemotaxis, only 5 genes were up-regulated *in vivo*, while 25 genes were down-regulated (Supplementary table S7). This pattern is in accordance with the loss of motility of in the *in vivo* bacteria (Fig. 1l, Supplementary fig. S3).

[*Type VI secretion system*]

Another envelope complex that is strikingly down-regulated in the *in vivo* symbiotic cells is the Type VI secretion system (T6SS). *Burkholderia insecticola* strain RPE64 possesses two T6SS-encoding loci, one on the chromosome 1 and one on chromosome 3, both of which were strongly suppressed in the symbiotic organ (Supplementary table S7). Two types of T6SS have been described: the first type is involved in antagonism between bacteria by injecting toxic effectors in competing bacteria, while the second type is used for interacting with eukaryotic cells and therefore it is a virulence factor (Russell et al., 2014). Although the biological function of the symbiont T6SSs remains unclear, the *in vivo* down-regulation strongly suggests that the T6SSs do not play a considerable function in the gut symbiotic organ. The down-regulation in the crypts

however is not incompatible with an active role of the T6SS earlier in the infection process, for example during the passage of the bacteria in the anterior midgut or the constricted region (Ohbayashi et al., 2015) where competition with other bacteria and/or active interaction with host cells might be important.

[*Polyhydroxylalkane (PHA) biosynthesis*]

PHA is a carbon and energy storage compound and a number of bacteria generate it from lipid and sugar (Anderson and Dawes, 1990). Previous genetic studies have revealed that the *Burkholderia* symbiont needs PHAs to stably colonize the midgut crypts (Kim et al., 2013; Jang et al., 2017). Although *in vivo* up-regulation of the PHA biosynthesis pathway has been detected at the protein level (Kim et al., 2013), the RNA-seq analysis demonstrated that this pathway is not up-regulated *in vivo* at the transcription level (Supplementary Table S7).

[*Other metabolism*]

We here point out three notable gene clusters that are highly up-regulated *in vivo* (Supplementary Table S7), although their biological functions are not clearly annotated. The first cluster (ACDS04980-ACDS05050) consists of membrane-associated proteins such as K^--sensing two-component system, protease Do and OmpA/MotB, probably involved in a bacterial stresss response. The second cluster (BCDS01210- BCDS01720) consists of pyrroloquinoline quinone (PQQ) synthase, PQQ-dependent dehydrogenase and putative methanol dehydrogenase, which may be involved in C1 metabolism. The

third cluster (BCDS13730-BCDS14040) consists of genes probably involved in serine metabolism, such as serine deaminase and threonine dehydrogenase. In addition to the three clusters, a number of functionaly-unknown ABC transporters showed *in vivo* specific up-regulation suggesting that the host supplies the symbionts with a variety of specific compounds (Supplementary Table S7).

Supplementary References

Anderson AJ, Dawes EA. Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol Rev.* 1990; **54**: 450-472.

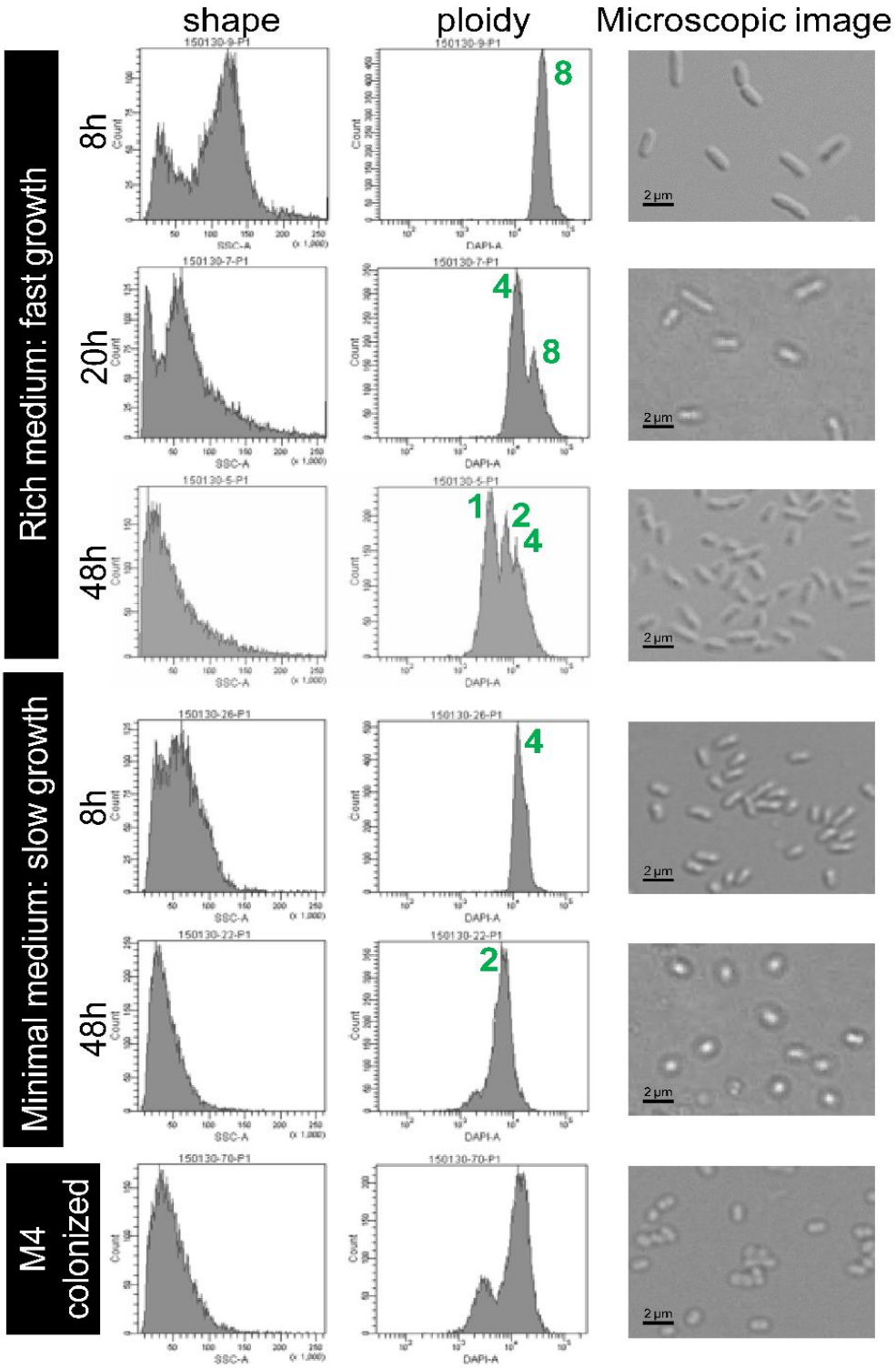
Kim JK, Won YJ, Nikoh N, Nakayama H, Han SH, Kikuchi Y *et al*. Polyester synthesis genes associated with stress resistance are involved in an insect-bacterium symbiosis. *Proc Natl Acad Sci USA.* 2013; **110**: E2381-2389.

Jang SH, Jang HA, Lee J, Kim JU, Lee SA, Park KE *et al*. PhaR, a negative regulator of PhaP, modulates the colonization of a *Burkholderia* gut symbiont in the midgut of the host insect, *Riptortus pedestris*. *Appl Environ Microbiol.* 2017; **83**: e00459-17.

Ohbayashi T, Takeshita K, Kitagawa W, Nikoh N, Koga R, Meng XY *et al*. Insect's intestinal organ for symbiont sorting. *Proc Natl Acad Sci USA.* 2015; **112**: E5179–E5188.

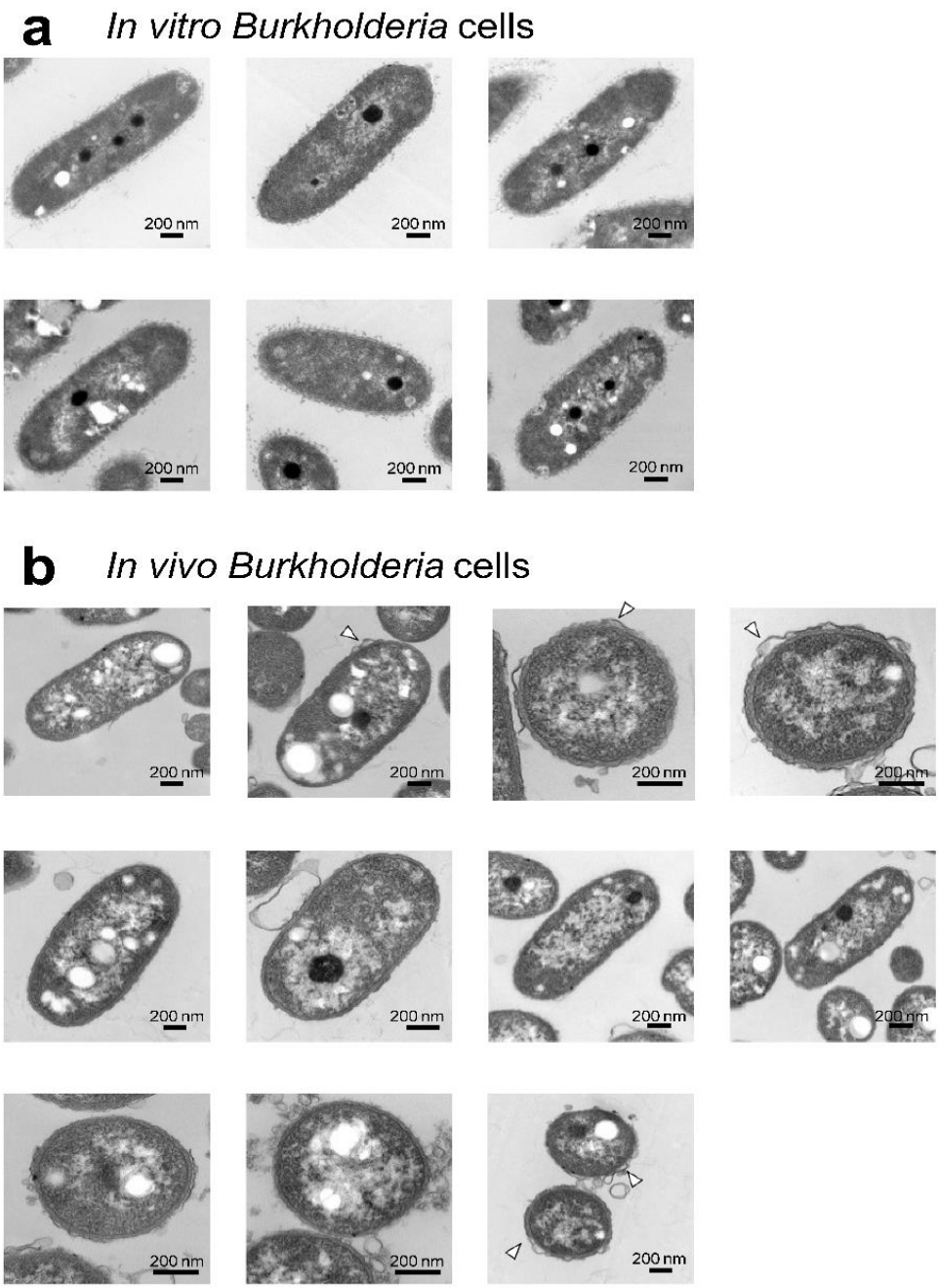
Russell AB, Peterson SB, Mougous JD. Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol.* 2014; **12**: 137-148.

Willis L, Huang KC. Sizing up the bacterial cell cycle. *Nat Rev Microbiol.* 2017; **15**: 606.

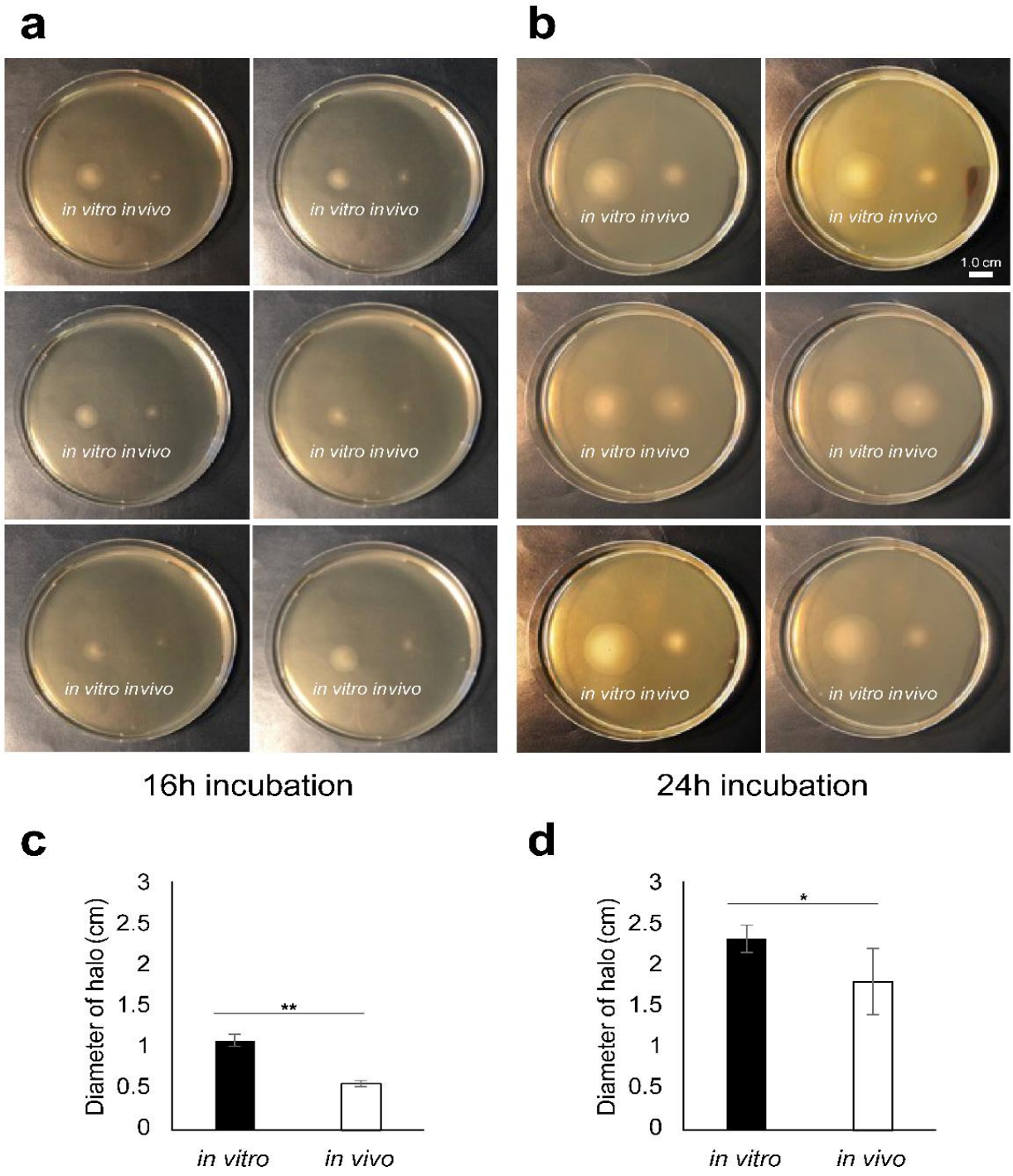


Supplementary Fig. S1 The cell size and DNA contents of the *Burkholderia* symbiont in different growth phases and culture medium.

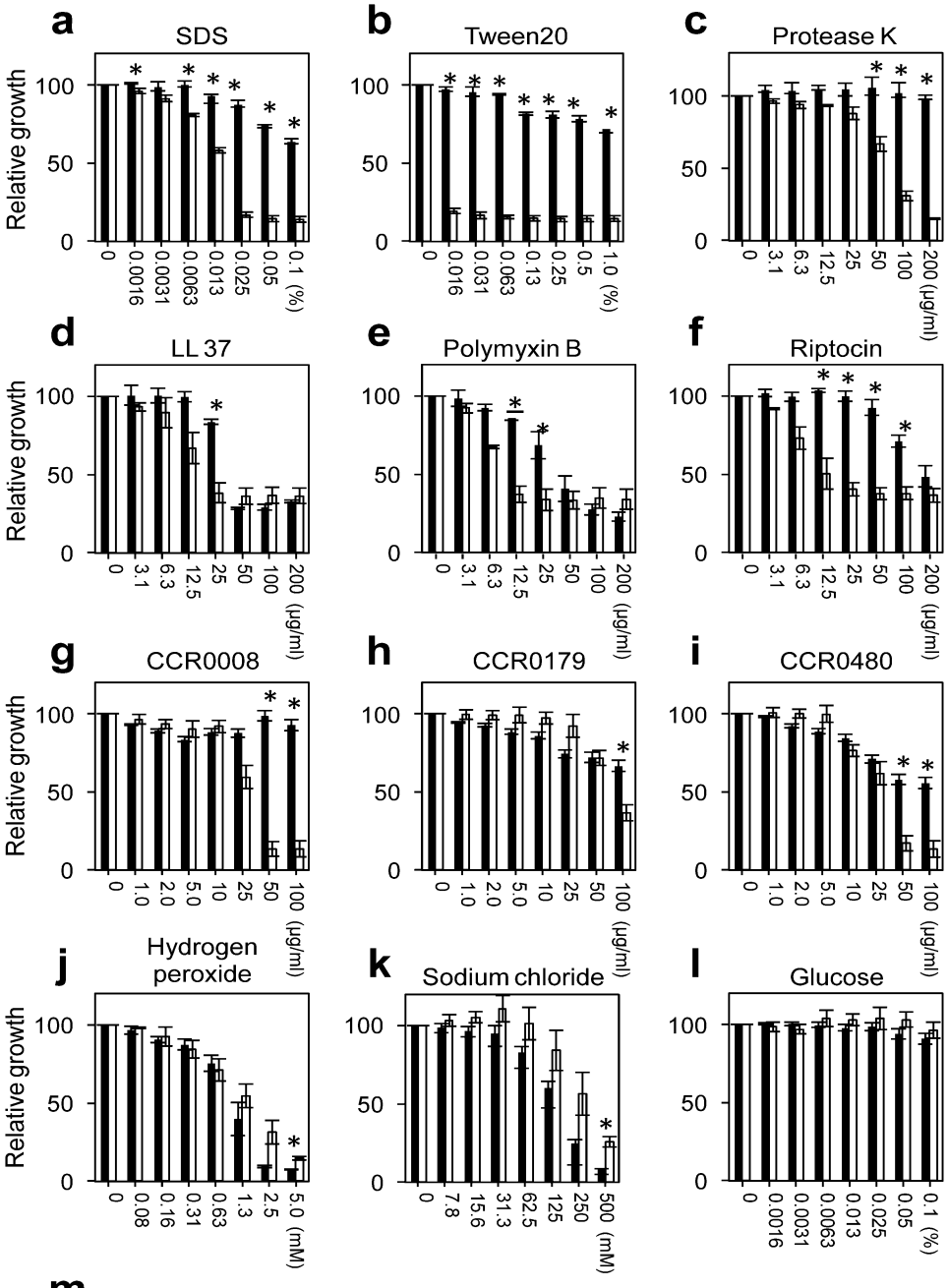
Flow cytometry analysis of the cell size by light scatter and the DNA content by DAPI staining and the microscopic observation of the *in vivo* and *in vitro* bacteria. The *in vitro* bacteria were grown for 8 h, 20 h and 48 h in the rich YG medium and 8 h and 48 h in the minimal medium.



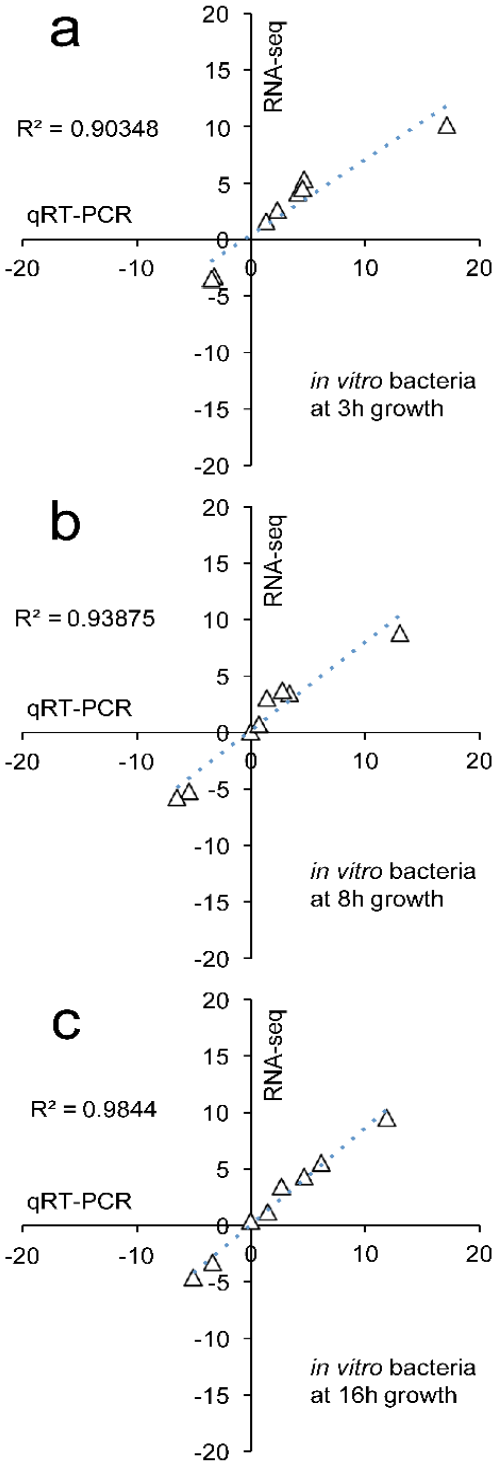
Supplementary Fig. S2 Transmission Electron Microscopic images of *in vivo* and *in vitro Burkholderia* cells
a No cells with a breb out of 6 *in vitro* cells. **b** The five positive cells with a breb out of 12 *in vivo* cells. Arrowhead shows a breb.



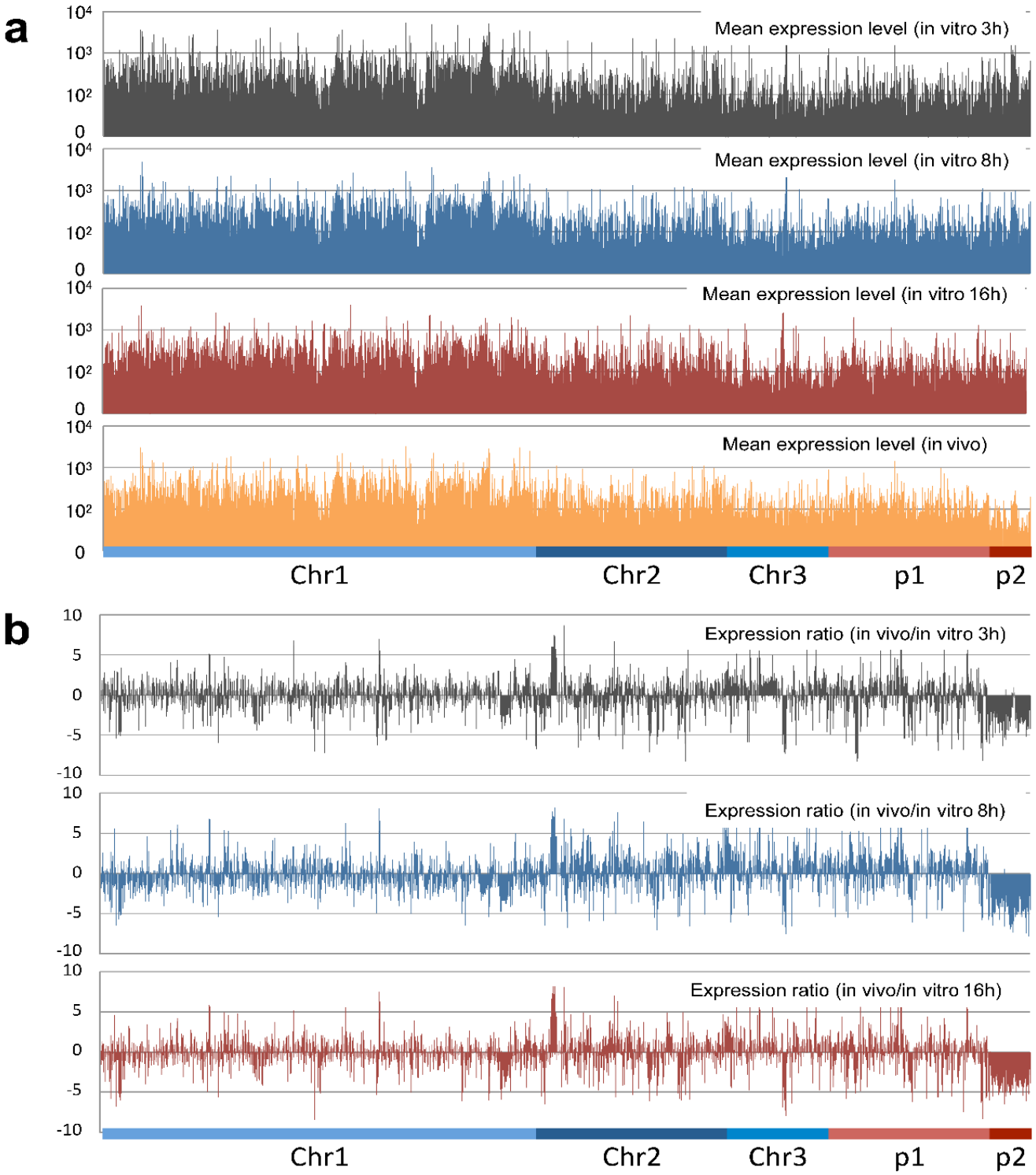
Supplementary Fig. S3 Flagellar motility of *in vivo* and *in vitro* bacteria
a-b Images of Flagellar motility of *in vivo* and *in vitro* bacteria in 0.4% YG soft-agar plate at **a** 16 and **b** 24 hours incubation. **c-d** Average of diameter of halo in *in vivo* (solid bar) and *in vitro* (open bar) bacteria at **c** 16 and **d** 24 hours incubation. Error bar indicates standard deviation (n=6). Spastically significant difference between *in vivo* and *in vitro* bacteria were analyzed by Student's *t*-test: ** $p<0.01$, * $p<0.05$.



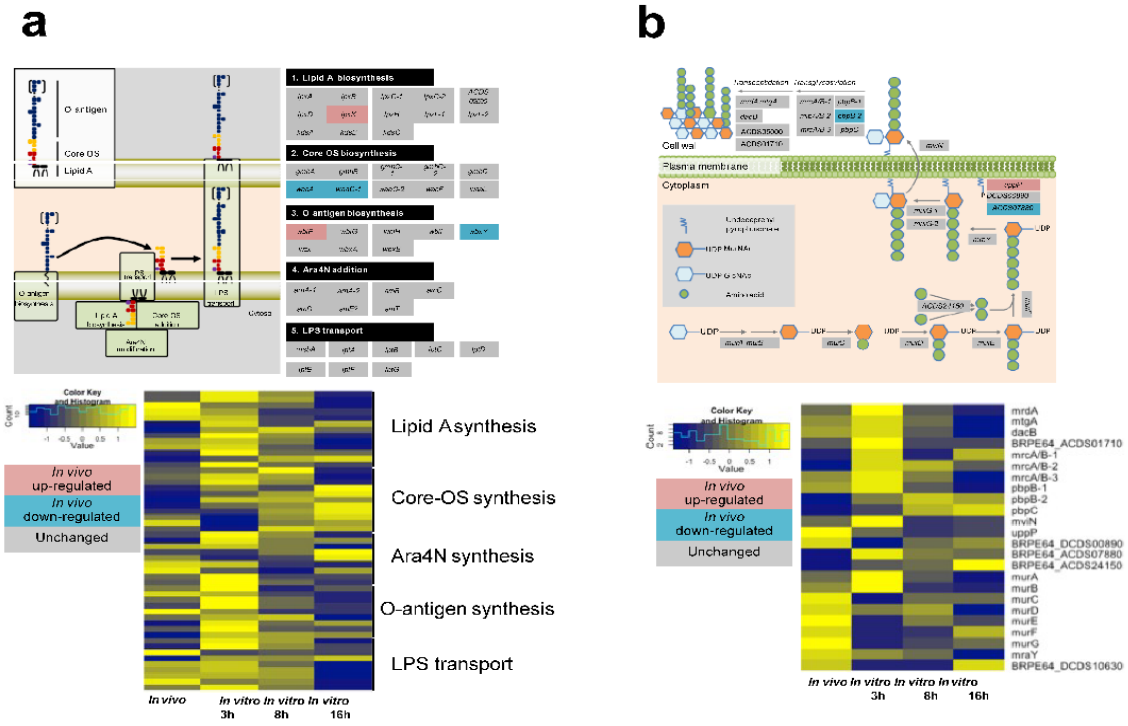
Supplementary Fig. S4 Comparison of stress sensitivities between *in vivo* and *in vitro* bacteria in the liquid media. The *in vivo* and *in vitro* *Burkholderia* symbiont cells were exposed to **a** SDS, **b** tween 20, **c** protease K, **d** LL-37, **e** polymyxin B, **f** riptocin, **g** CCR0008, **h** CCR0179, **i** CCR0480, **j** hydrogen peroxide, **k** sodium chloride, **l** glucose, and **m** ethanol. The relative growth (%) was normalized by untreated bacterial growth set to 100 %. Mean ± SD (n = 3) is shown in *in vitro* (solid bar) and *in vivo* cells (open bar). Statistically significant differences between *in vivo* and *in vitro* symbiont cells were analyzed by t-test with Bonferroni correction: $P < 0.001$.



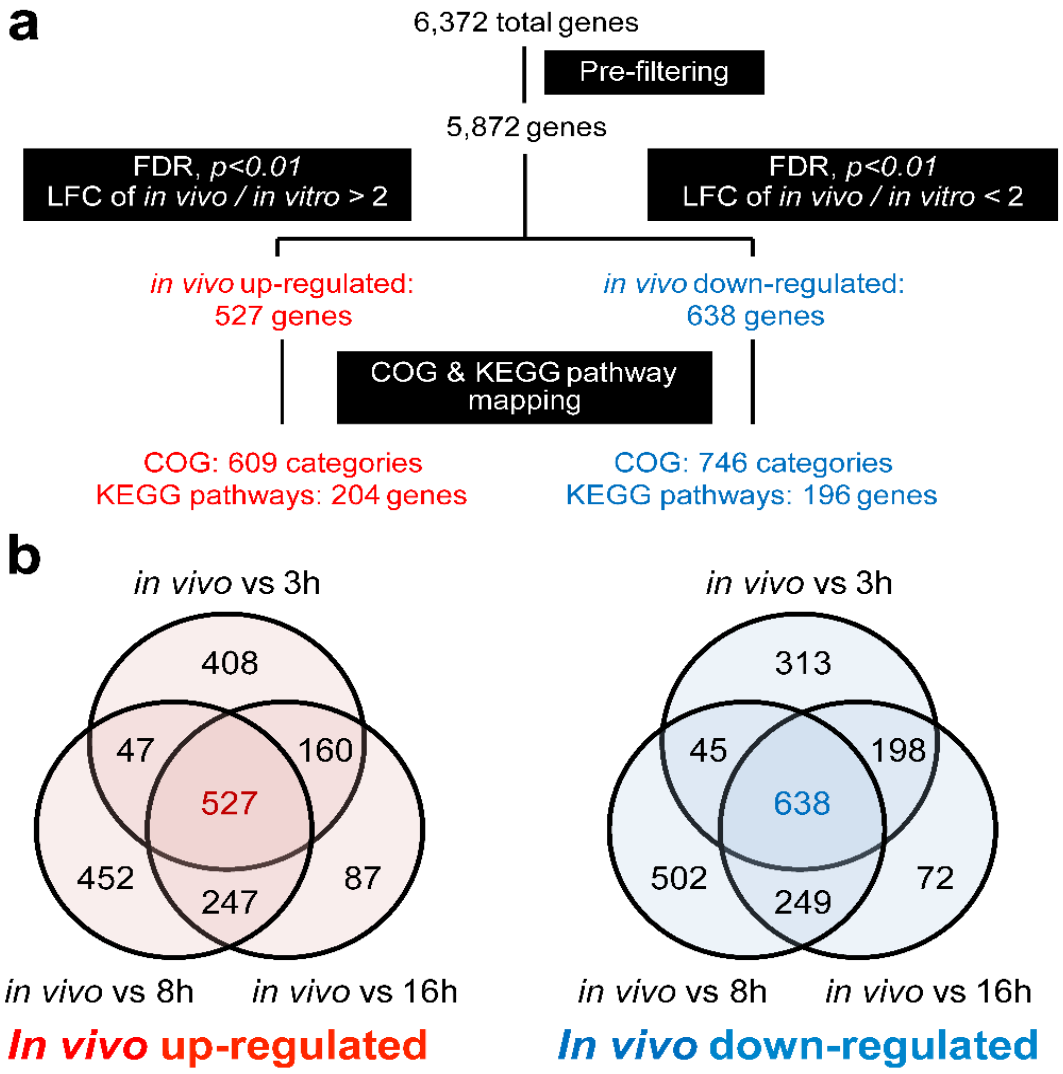
Supplementary Fig. S5 Validation of RNA-seq data by qRT-PCR.
a-c The relative gene expression levels of 8 randomly selected genes were compared between RNA-seq and qRT-PCR data. The qRT-PCR data were processed with the $\Delta\Delta C_T$ method. The relative expression levels were calculated as the *in vivo* per *in vitro* value. Comparisons are shown for *in vitro* bacteria at 3h growth (a), 8h growth (b) and 16h growth (c). Correlation values (R^2) of the fold change between RNA-seq (y-axis) and qRT-PCR (x-axis) are shown.



Supplementary Fig. S6 The global gene expression profile of *the in vivo and in vitro Burkholderia insecticola* RPE75.
a The mean expression level of 5,872 genes in *in vitro* 3h, 8h and 16h culture and *in vivo* bacteria. **b** The expression ratio of the RNA-seq data in *in vivo per in vitro* 3h, 8h and 16h culture bacteria. The genes are ordered as in the genome and colors of x-axis indicate the five genome components of *Burkholderia insecticola* strain RPE64 in the order chromosome 1, chromosome 2, chromosome 3, plasmid 1, plasmid 2.



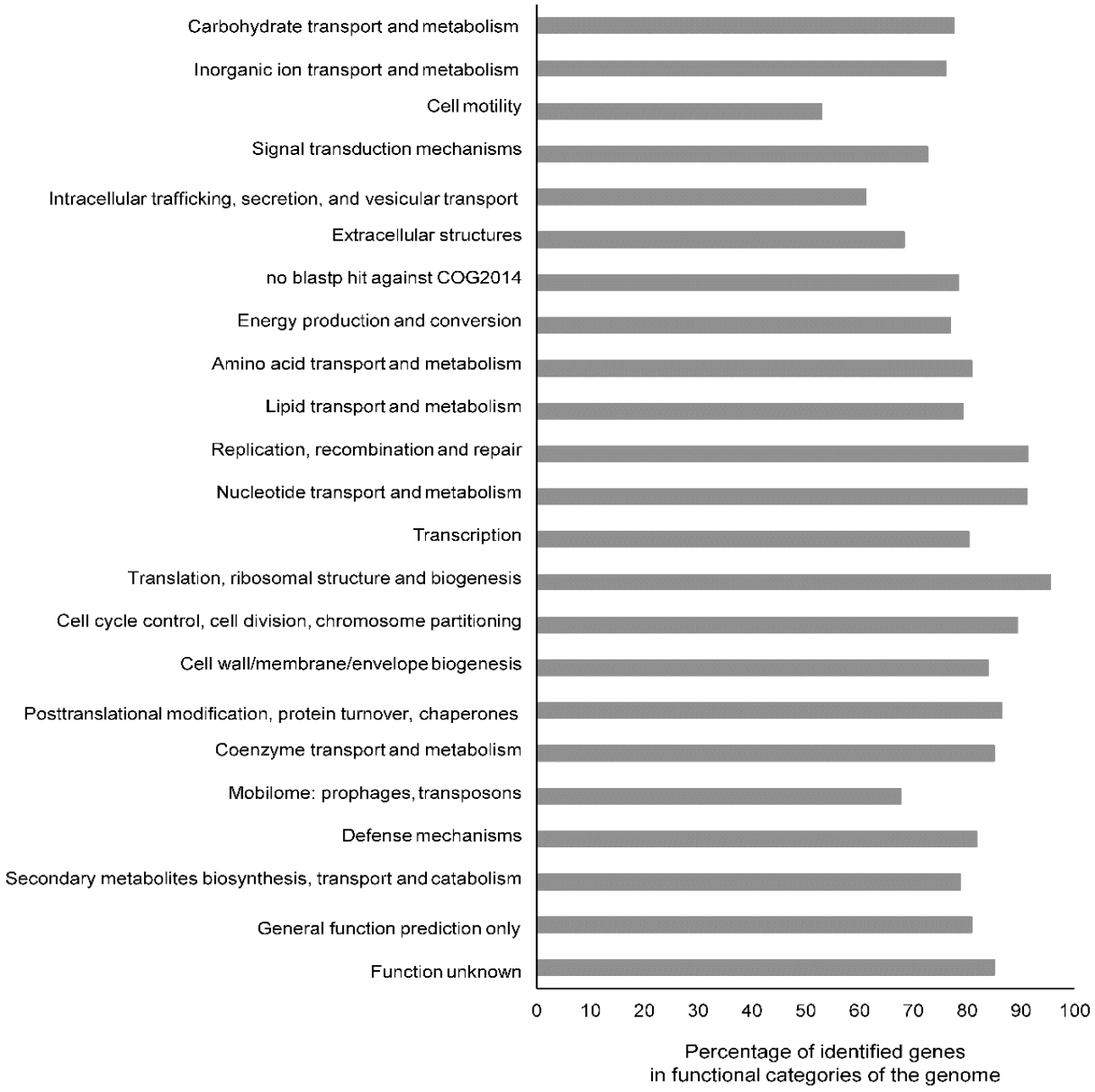
Supplementary Fig. S7 Gene expression profile of surface structure in *in vivo* and *in vitro* cells.
a The expression profile of 54 LPS biosynthesis genes in *in vivo* bacteria. LPS is the major constituent of the outer membrane in gram-negative bacteria. Its biosynthesis is composed of five main sections: lipid A biosynthesis, the core oligosaccharide (OS) biosynthesis, the O-antigen biosynthesis, 4-amino-4-deoxy-L-arabinose (Ara4N) addition, and LPS transport (Raetz et al., 2007; Maldonado et al., 2016). **b** The expression profile of 24 genes in the peptidoglycan biosynthesis pathway. Arrows indicate metabolic direction, and the color shows gene expression levels in RNA-seq: red for *in vivo* up-regulated; blue for *in vivo* down-regulated; gray for unchanged expression level. Abbreviations: GlcNAc, N-acetylglucosamine; MurNAc, N-Acetylmuramic acid; UDP, uridine diphosphate.



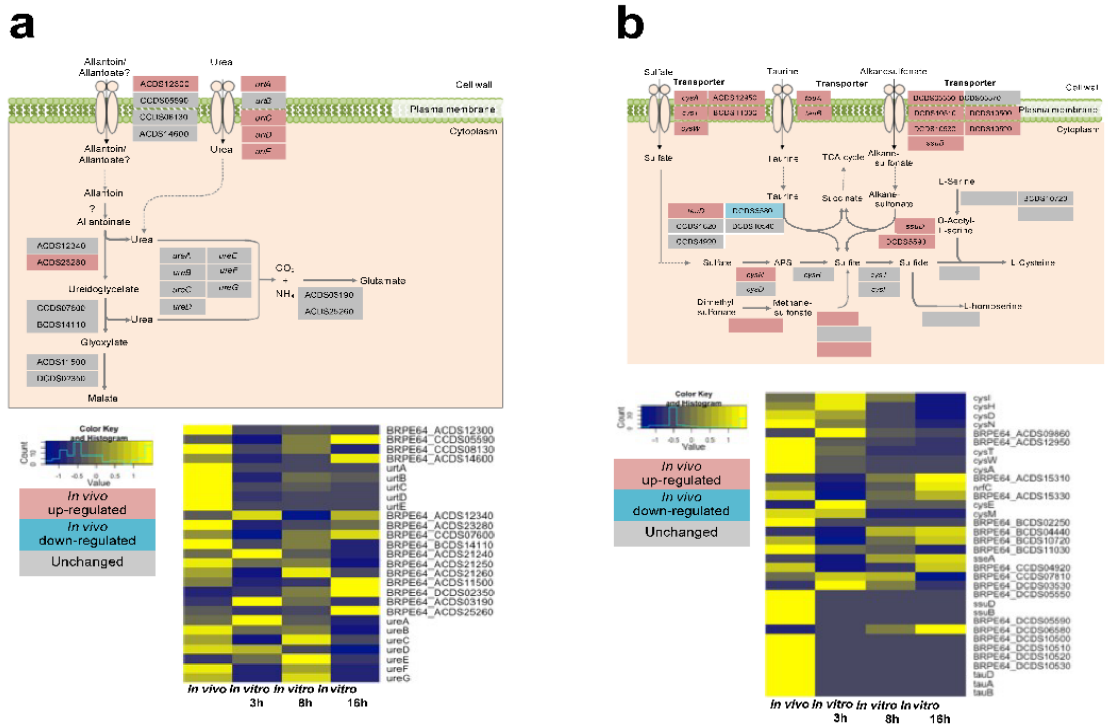
c

	Up-regulated specifically in vivo	Down-regulated specifically in vivo	Un-changed	Total gene numbers
Chromosome1	166	234	2334	2734
Chromosome2	165	99	998	1262
Chromosome3	73	32	483	588
Plasmid1	123	83	813	1019
Plasmid2	0	190	79	269
Total	527	638	4707	5872

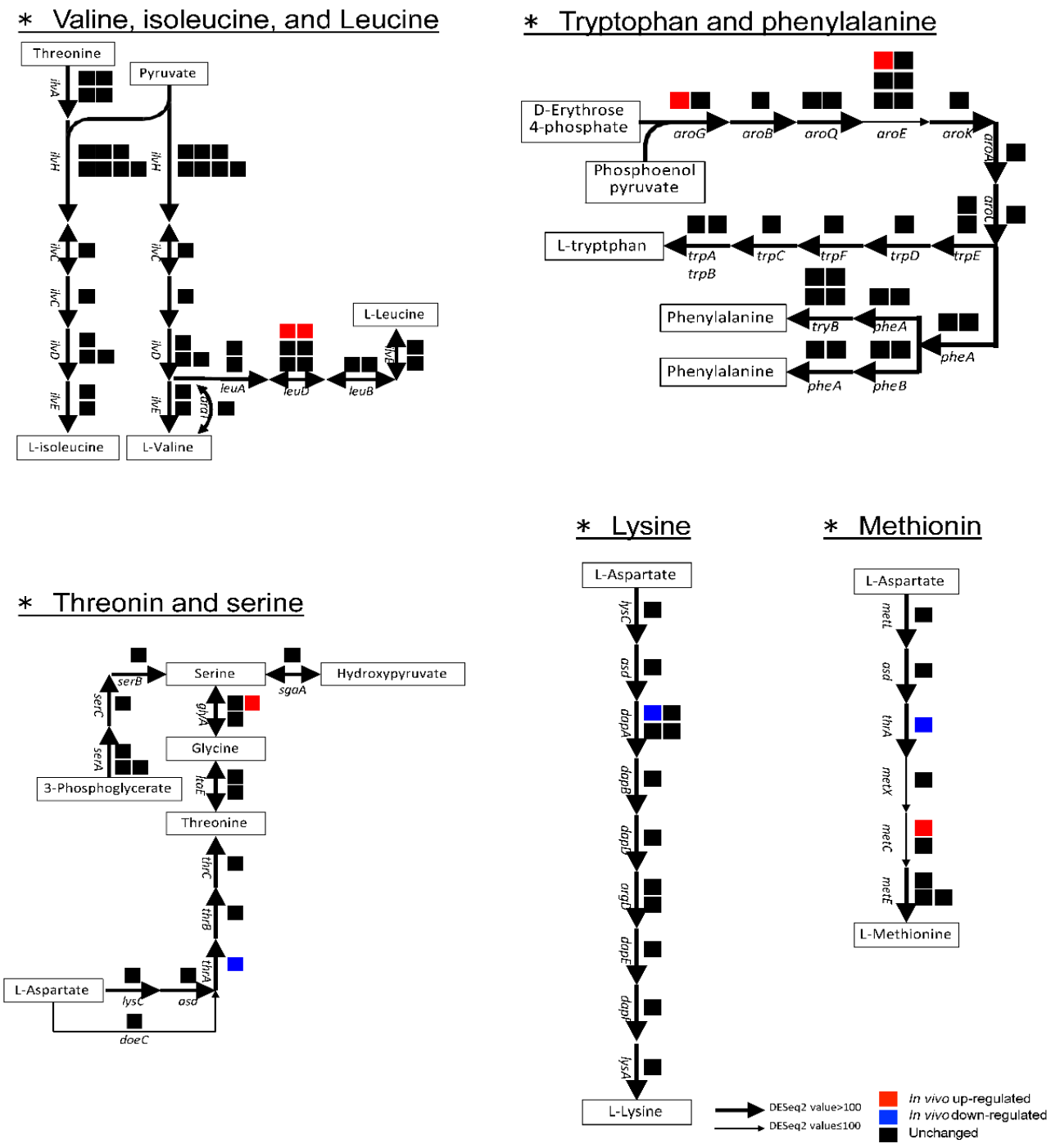
Supplementary Fig. S8 The determination process of *in vivo* up-regulated and down-regulated genes in RNA-seq
a The determination of *in vivo* up-regulated and down-regulated genes from the total gene set and classification in COG categories and KEGG metabolic pathways. **b** Venn diagram showing gene numbers of significant differential expression (FDR, *p*-value<0.01) and LFC of *in vivo/in vitro*>1 (left) and <1 (right) between three comparison, *in vivo/in vitro*-3h, *in vivo/in vitro*-8h and *in vivo/in vitro*-16h. Overlap of three comparison shows number of *in vivo* up-regulated genes (left) and down-regulated genes (right).



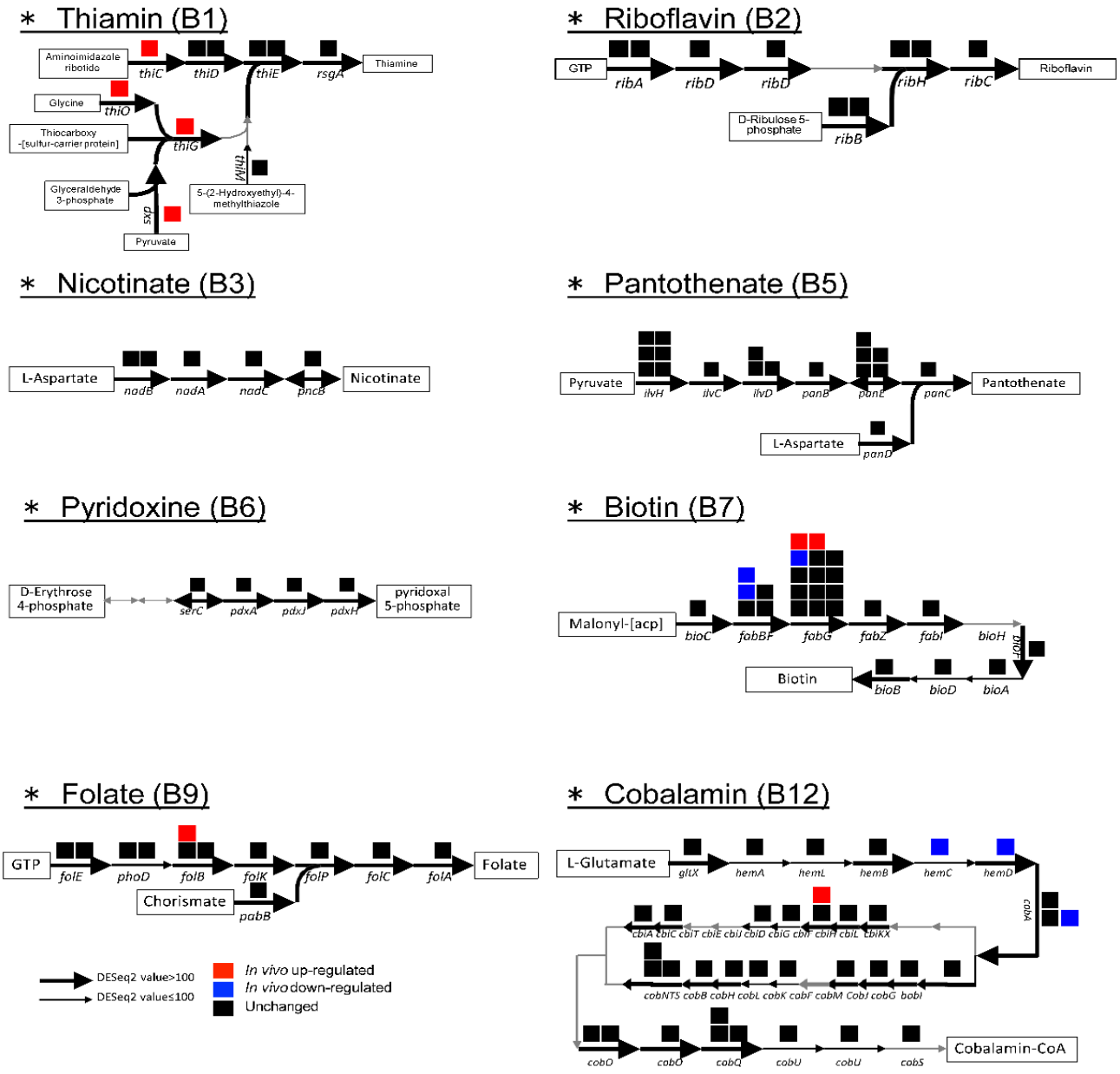
Supplementary Fig. S9. COG functional classification of *in vivo* un-changed genes.



Supplementary Fig. S10 Highlight of *in vivo* up-regulated pathways.
a Allantoin and urea metabolism. The expression profile of 15 genes in the allantoin metabolic pathway. This pathway was inferred from Navone et al., 2015. **b** Sulfur metabolism. The expression profile of 32 genes in the sulfur metabolic pathway. Arrows indicate metabolic direction, and the color shows gene expression levels in RNA-seq: red for *in vivo* up-regulated; blue for *in vivo* down-regulated; gray for unchanged expression level. Abbreviations: APS, adenosine 5'-phosphosulfate.



Supplementary Fig. S11 *In vivo* gene expressions in biosynthesis pathways of 9 essential amino acids.
The 177 gene expression profile of 9 essential amino acids biosynthesis in *in vivo* bacteria. These pathways were referred from KEGG metabolic pathway database. Arrows indicate metabolic direction, and sizes indicate that large and small are a high and low expression gene (FPKM>100, FPKM≤100), respectively. Gray arrows indicate no gene annotated in KEGG database. Each block indicates annotated genes in KEGG database. the color shows gene expression levels in RNA-seq: red for *in vivo* up-regulated; blue for *in vivo* down-regulated; gray for unchanged expression level.



Supplementary Fig. S12 *In vivo* gene expressions in B-vitamins biosynthesis pathways.
The 208 gene expression profile of 8 B-vitamins biosynthesis in *in vivo* bacteria. These pathways were referred from KEGG metabolic pathway database. Arrows, blocks, and colors are the same as Supplementary fig.4, and gray arrows are no annotated genes in KEGG database. Abbreviations: GTP, guanosine diphosphate; CoA, Coenzyme A; [acp], [acyl-carrier protein].

Titre : Analyse comparative des mécanismes de différenciation des bactéroïdes au cours des symbioses *Aeschynomene-Bradyrhizobium*

Mots clés : Symbiose fixatrice d'azote, bactéroïdes, polyploïdie, transcriptomique, protéomique, métabolomique, efficacité symbiotique

Résumé : En cas de carence azotée, les légumineuses sont capables de mettre en place une symbiose avec des bactéries du sol fixatrices d'azote appelées rhizobia. Cette symbiose a lieu dans un organe appelé nodosité où les bactéries sont endocytées et appelées bactéroïdes. Certains clades de légumineuses imposent un processus de différenciation à leurs bactéroïdes qui agrandissent considérablement et deviennent polyploïdes, menant à des morphotypes bactériens allongés ou sphériques. Au cours de cette thèse, j'ai étudié la différenciation des bactéroïdes de *Bradyrhizobium* spp. en association avec *Aeschynomene* spp.. Les bactéroïdes de ces plantes présentent des degrés de différenciation distincts qui dépendent de l'espèce hôte. Mes données suggèrent que les bactéroïdes les plus différenciés sont aussi les plus efficaces. J'ai cherché à savoir quels facteurs procaryotes pourraient être impliqués dans les adaptations des bactéroïdes au processus de différenciation et à leurs divers hôtes, le tout en lien avec cette différence d'efficacité symbiotique au travers d'approches globales sans *a priori* de type -omiques. Les conditions considérées sont des bactéroïdes de différents morphotypes et des cultures libres de référence. Les fonctions activées en conditions symbiotiques ont été identifiées, ainsi que les gènes spécifiques d'un hôte donné. Des analyses fonctionnelles des gènes d'intérêt ont également été menées. Les mutants bactériens n'ont toutefois pas présenté de phénotype symbiotique drastique, montrant ainsi l'existence de réseaux de gènes complexes menant à la résilience des génomes de rhizobia.

Title : Comparative analysis of bacteroid differentiation mechanisms in *Aeschynomene-Bradyrhizobium* symbioses

Keywords : Nitrogen-fixing symbiosis, bacteroids, polyploidy, transcriptomics, proteomics, metabolomics, symbiotic efficiency

Abstract: In case of nitrogen starvation, legume plants establish a symbiotic interaction with nitrogen-fixing soil bacteria called rhizobia. This interaction takes place in nodules where the symbionts are internalized and become bacteroids. Some legume clades also impose a differentiation process onto the bacteroids which become enlarged and polyploid, leading to elongated or spherical morphotypes. During my PhD work, I have studied bacteroid differentiation of *Bradyrhizobium* species in association with *Aeschynomene* spp.. These bacteroids display distinct differentiation levels depending on the plant host, and my analyses suggest that the most differentiated ones are also the most efficient. I investigated the bacterial factors potentially involved in the adaptations to differentiation and host-specificity, and related to the higher efficiency of the most differentiated bacteroids using global-omics approaches without *a priori*. The analyzed conditions were bacteroids of distinct morphotypes and free-living reference cultures. Activated functions under symbiotic conditions were identified, as well as host-specific ones. Functional analyses were performed on genes of interest. However, the bacterial mutants did not display drastic symbiotic phenotypes, showing the existence of complex gene networks leading to high resilience of rhizobial genomes.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France