THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Biologie Des Interactions

École doctorale GAIA – Mécanismes des Interactions parasitaires pathogènes et symbiotiques

LSTM – Laboratoire des Symbioses Tropicales et Méditerranéennes
PHIM – Plant Health Institute of Montpellier

Caractérisation fonctionnelle d'une nouvelle classe d'effecteurs de type III capables de déclencher l'organogenèse nodulaire chez les légumineuses.

Présentée par Alicia CAMUEL
Le 19 Décembre 2023

Sous la direction du Dr. Eric GIRAUD

Devant le jury composé de

Pr. Sofie GOORMACHTIG, Professeure des Universités, Université de Ghent-VIB-UGENT	Rapporteure
Pr. Matthieu ARLAT, Professeur des Universités, Université de Toulouse-LIPME	Rapporteur
Pr. Brigitte BRUNEL, Professeure des Universités, Montpellier SupAgro-LSTM	Présidente du jury
Dr. Delphine CAPELA, Directrice de recherche, CNRS-LIPME	Examinatrice
Dr. Eric GIRAUD, Directeur de recherche, IRD-PHIM	Directeur de thèse

UNIVERSITÉ DE MONTPELLIER

“Dans la vie, les jeux sont donnés, mais avec un jeu donné, chacun peut faire une partie différente.”

Johann Wolfgang von Goethe

Remerciements

Cela fait maintenant trois ans que j'ai ouvert la page de cette aventure qu'est la thèse et il faut maintenant la refermer. J'ai vécu et appris tellement de choses durant ces trois années, et, en même temps, elles sont passées à une telle vitesse que j'ai du mal à réaliser qu'aujourd'hui c'est la dernière page que j'écris. Et, comme toutes les belles histoires, celle-ci est faite de rencontres qui l'ont rendue si riche en émotions, en découvertes et en rebondissements. Chaque personne que j'ai croisée a eu son importance d'une manière ou d'une autre, et je tiens donc dans les paragraphes suivants à les remercier.

Dans un premier temps, je tiens à remercier les membres de mon jury de thèse, qui ont accepté d'évaluer mon travail et de faire le déplacement pour assister à ma soutenance : mes rapporteurs **Sofie Goormachtig** et **Matthieu Arlat** ainsi que mes examinateurs **Brigitte Brunel** et **Delphine Capela**. Un grand merci également aux membres de mes comités de thèse, **Peter Mergaert**, **Laurent Deslandes** et **Guilhem Desbrosses**. Vous m'avez suivie et conseillée durant toute cette période de thèse et j'ai eu énormément de plaisir à échanger avec vous. Merci à tous nos collaborateurs qui ont contribué énormément à ce travail de thèse et qui m'ont fait prendre conscience de l'importance et de la richesse des collaborations scientifiques. Parmi eux je remercie à nouveau Peter et Laurent, mais également **Haiwei Luo**, **Fazal Haq**, **Mélanie Carcagno**, **Lu Ling**, **Shin Okazaki** et **Safirah Tasa Nerves Ratu**. Je tiens à remercier aussi nos collaborateurs thaïlandais qui ont travaillé à nos côtés et m'ont accueillie dans leur magnifique pays : **Neung Teaumroong**, **Panlada Tittabutr**, **Komonluck Teamtisong**, **Pongdet Piromyou**, **Jenjira Wongdee**, **Pongpan Songwattana** et **Bidou Aphaiso**.

C'est avec beaucoup d'émotion que je tiens à remercier mon directeur de thèse **Eric Giraud**. C'est à tes côtés que j'ai pris conscience de mon désir de faire une thèse, et, avec le recul, je pense que je n'aurais pas pu la réaliser auprès de quelqu'un d'autre ou ailleurs qu'au LSTM. Je n'aurais pas pu rêver mieux qu'une thèse sur ce modèle extraordinaire qu'est *Bradyrhizobium*/*Aeschynomene*, qui cache encore plein de surprises et qui m'avait initialement attiré pour mon stage de M1. Mais, outre la curiosité scientifique et la passion pour le sujet que tu as su me transmettre, je tiens à te remercier aussi pour la personne que tu es. Tu as toujours été à l'écoute en réponse à mes innombrables "toc toc" frappés à ta porte, toujours patient pour prendre le temps de m'expliquer et de me conseiller au mieux. En plus d'être un excellent scientifique, c'est ton enthousiasme, ton humour, ta gentillesse et ta bienveillance qui font de toi une personne que je n'oublierai jamais. Tu m'as permis de réaliser des choses que je n'aurais jamais pu imaginer faire dans ma vie (en plus de cette thèse !). Je pense notamment à mes voyages en Thaïlande et aux États-Unis qui m'ont fait grandir intérieurement et ont imprégné ma

mémoire de merveilleux souvenirs et de merveilleuses rencontres. Et, en parlant de souvenirs, je n'en ai aucun d'un seul mauvais moment passé à tes côtés ni aucun conflit quel qu'il soit, et, mon papa pourra te le dire, j'ai pourtant une fâcheuse habitude à tenir tête lorsque je suis en désaccord avec quelqu'un. Tout cela pour dire à quel point notre duo a bien fonctionné et à quel point il va aussi me manquer.

Je ne peux finalement pas remercier Eric sans aussi penser à **Albin Teulet**. Initialement étudiant et doctorant avant moi, tu as été tout comme Eric un élément déclencheur de mon choix de réaliser cette thèse. Je sais que beaucoup de personnes vont sourire en lisant ce paragraphe, mais tu as été un réel mentor pour moi, tu m'as appris tant de choses, qu'elles soient techniques ou théoriques, que je te dois aussi beaucoup pour la personne que je suis aujourd'hui. J'admire le scientifique que tu es devenu et j'apprécie énormément la personne que tu es. J'espère que nos chemins continueront de se croiser dans le futur, et, qui sait, peut-être qu'un jour nous travaillerons ensemble ?

J'ai une grande pensée aussi pour mes acolytes et amis de thèse : Natasha et Kenji. **Natasha**, j'ai été tellement heureuse de partager mon bureau jusqu'à très récemment avec toi. Tu as sans doute été la plus belle rencontre de ma thèse, tu es devenue une amie irremplaçable. Tu me comprends sur tellement de choses, tu es de merveilleux conseils et tellement à l'écoute. J'espère pouvoir t'aider comme toi tu m'as aidée quand tu auras besoin d'une épaule sur laquelle te reposer. Tu es une personne formidable et surtout la plus belle des Brésiliennes ! J'espère qu'il y aura d'autres chemins sur lesquels nous marcherons ensemble après nos thèses respectives. **Kenji**, nous avons commencé notre thèse le même jour et je suis impressionnée par tout le travail que tu as accompli. Tu vas devenir un excellent scientifique, j'en suis certaine. Et, malgré cette charge de travail, tu as toujours pris du temps pour venir me voir, discuter avec moi quand j'en avais besoin et être un ami généreux sans oublier tes talents culinaires ! Je te souhaite tout le meilleur pour la suite.

Réaliser une thèse implique aussi de travailler au sein d'une équipe et aux côtés d'autres personnes, qui, ensemble, font vivre le laboratoire. C'est pourquoi je tiens tout particulièrement à remercier ceux qui ont rendu cette thèse si merveilleuse à vivre.

Merci à toute l'équipe MSLT :

Jean-François Arrighi, notre super nouveau chef d'équipe pour qui j'ai beaucoup de respect et d'admiration. Tu es le puits de connaissance sur les *Aeschynomene* que tout le monde, je suis sûre, jalouse. Je me souviendrai toujours des mots d'encouragements que tu m'avais transmis suite à ma présentation devant l'équipe durant mon stage de M1 et qui m'ont donné confiance en moi dans la perspective d'une thèse. **Marjorie Pervent**, j'ai appris à tellement mieux te connaître durant toute ma thèse. Tu es

quelqu'un de merveilleux et tu donnes tellement de ta personne. Ton aide tout comme ton sourire m'ont été très précieux et je t'en remercie. **Frédéric Gressent**, avec qui, avouons-le, j'ai plus papoté que travaillé ! Tu réponds néanmoins toujours présent quand on a besoin de toi et ta bonne humeur est tellement agréable au quotidien, merci. **Nico Nouwen**, qui pétille toujours de diverses idées et hypothèses scientifiques ! Et qui connaît aussi les meilleures adresses pour trouver du super gouda ! Merci d'être un noyau fort de notre équipe sur qui on peut toujours compter. **Clémence Chaintreuil**, merci pour tes mots d'encouragement que tu n'as cessé de m'adresser durant toute la rédaction de cette thèse et même avant. Enfin, un merci tout particulier à ma nouvelle voisine de bureau, **Maëlle Rios**, qui apporte beaucoup à notre équipe et qui nous apparaît de plus en plus indispensable. J'ai trouvé en Maëlle une acolyte pour partager ma passion pour le sport ! Allez les Bleus !!

Merci à tout le LSTM :

Bien que l'histoire du LSTM se terminera bientôt, je n'oublierai jamais que c'est avec vous tous que j'ai pu m'épanouir pendant ces trois ans. Merci d'abord à **Robin Duponnois**, pour son accueil au sein de l'unité et pour tous les échanges que nous avons pu avoir, teintés souvent d'humour mais aussi de perspicacité scientifique. Cet accueil a été également une réussite grâce à **Cathy Mortier** et **Nathalie Claudé**. J'ai tout particulièrement appris à mieux te connaître, Nathalie, durant ces quelques semaines de confinement passées au labo au début de ma thèse, où j'avais pris la place du grand chef et où nos bureaux se trouvaient adjacents. Tu es une personne merveilleuse, sincère et toujours là si besoin. Tu me manques déjà. Merci aussi à **Yves,** le roi de la brocante, **Sanaa** aux bérets magnifiques, **Ezekiel** le grand sportif, **Antoine** le nageur en devenir, **Christine** la fée du thé glacé, **Marc** le roi des jeux de mots, **Julie** la nouvelle maman et **Naïma** notre maman à tous, merci pour les pauses café et les breaks périodiques du jeudi que nous avons partagés, tout comme les moments passés au laboratoire à se soutenir et à s'aider mutuellement. Ces moments de partage sont aussi ce qui fait la vie d'un labo et resteront de très beaux souvenirs.

Je tiens à remercier aussi l'UMR PHIM dont je fais maintenant partie et où les anciens membres de l'équipe MSLT se retrouvent aujourd'hui. Je pense particulièrement à **Fabienne Cartieaux** et **Djamel Gully** qui m'ont connue dès mon stage de M1 et m'ont apporté leur aide durant cette thèse.

Merci à tous les non-permanents :

Parmi vous, je me dois de commencer par le meilleur duo qui soit, un duo d'inséparables, où une taquinerie en cache toujours une autre : **Julien** et **Lucie**. Vous êtes devenus bien plus que des collègues pour moi. Vous m'avez toujours donné le sourire pour un oui ou pour un non, même si vous ne vous en êtes peut-être pas toujours aperçu, et cela m'a donné une force supplémentaire dans les moments plus compliqués. Vous apportez aussi beaucoup au labo, de par vos nombreuses initiatives telles que la pétanque ou le bowling. Et puis surtout, je vous dois de vous le dire : merci d'avoir supporté mes blagues et jeux de mots les plus farfelus. Le ridicule ne tue pas et il fait finalement surtout rire !

Merci aux expatriés du bâtiment C avec qui j'ai partagés de très bons moments même si je n'ai pas toujours pris le temps de traverser ce brin de pelouse qui vous sépare du bâtiment J. Je pense notamment à **Pierre-Alexandre**, **Alexandre S.**, **Bryan**, **Margaux**, **Alexandre G.** et **Johanna**. Merci aussi à **Hamza** de n'être jamais bien loin, que ce soit en Modus ou à son bureau !

Merci à tous ceux qui furent de passage mais que je n'oublie pas :

Mathilde, la reine du CRB, ma partenaire de TP, mon amie depuis la Licence et avec qui j'ai eu la chance de partager, pendant une partie de ma thèse, mon quotidien. Ta rigueur scientifique m'épatera toujours et je suis heureuse du chemin que nous avons toutes les deux accompli. Merci d'être toujours présente de près ou de loin. Merci aux anciens MSLT, je pense surtout à **Johan** et **Rémi**. Vous avez été des super compagnons de bureau, mais aussi des amis qui m'ont poussé à avoir confiance en moi, ce qui m'a permis de concrétiser du mieux que je pouvais ma thèse. Merci également à **Stéphane** et **Laurent**, vous êtes la sagesse incarnée parmi les non-permanents. Avec toujours des conseils avisés et des mots pour réconforter. Merci à **Nico**, **Fernando** et **Denis**, le trio de Bactolytix-Greenphage pour leur bonne humeur quotidienne et les rires qu'ils ont apportés. Enfin, merci à tous les étudiants de passage qui ont participé à la vie et à la bonne ambiance du labo, que j'ai pris plaisir à aider et à conseiller du mieux que je pouvais. Il y en a eu beaucoup et la liste serait longue, ainsi je préfère m'attarder sur ceux qui ont habité le bureau J106 : **Sarah**, **Camille**, **Colin** et plus particulièrement **Melissa** et **Vetea** avec qui j'ai pu partager mon quotidien, mon amour pour les sciences et qui ont contribué aussi à ce travail de thèse.

Réaliser une thèse m'a conduit à faire des rencontres qui n'étaient pas forcément prévues mais qui m'ont apporté beaucoup, autant humainement que scientifiquement. Je pense à notre collaboration avec l'Auvergne fièrement représentée par **Anne-Catherine Lehours** et **Antoine Vergne**. L'aventure des cabarets des oiseaux a été un immense plaisir à vivre à vos côtés. Un grand merci également à **Claire Mayer**, mon mentor durant cette dernière année de thèse à travers le programme de mentorat Femmes & Sciences. J'ai appris énormément à tes côtés en si peu de temps, nos échanges sont toujours d'une grande richesse et je ne vois jamais le temps passer. Je souhaite mentionner aussi **Joan Estevan**, qui m'a connue en 2018 et m'a fait découvrir le monde de la recherche durant mon stage de Licence. Merci pour ton encadrement, ta gentillesse et ta rigueur qui m'ont accompagné lors de mes premiers pas dans le monde scientifique.

Je remercie ma famille, **Maman**, **Papa**, ma **Mémé adorée** et **mon frère que j'aime tant**. Vous savez l'investissement personnel que cette thèse a représenté pour moi et vous avez toujours su m'épauler avec amour. Un grand merci à **Christine** aussi, tu fais partie de la famille et tu mérites tout autant ces remerciements pour tout ce que tu as fait pour moi depuis que je suis toute petite. Une grande pensée

aussi à mes oncles, tantes, cousins et cousines, qui m'ont toujours soutenue dans mes études ainsi que pour mon **Papi** qui aurait, je pense, été très fier. Je remercie aussi mes amis, qui se trouvent à plus ou moins grande distance mais qui pensent toujours à moi et avec qui des liens forts se sont créés au fil des années. Merci au trio du lycée **Margot**, **Eric** et **Thomas,** à mes comparses de prépa **Léa** et **Emma** et au trio de la fac **Mathilde**, **Mélusine** et **Maxime.** Votre amitié m'est très précieuse.

Merci à mon bébé chat, mon **Roméo** qui fait chavirer mon petit cœur. On entend souvent dire que les animaux ressentent les émotions, et tu as toujours ressenti les miennes pour m'apporter ton amour et ton réconfort. Oui je suis gaga de mon chat, mais je vous garantis que vous le seriez aussi si vous le connaissiez.

Je terminerais ces remerciements par toi, **Quentin**. Je n'ai pas les mots pour te dire à quel point je te suis reconnaissante pour tout ce que tu m'as apporté depuis notre rencontre et pour ton soutien indéfectible au quotidien. Rien n'a été facile, mais nous avons partagé des moments forts qui nous ont permis à tous les deux d'avancer. Cette thèse ne serait pas la même sans toi, et je ne l'oublierai jamais. Mille fois merci, pour tout.

Enfin, lecteur, lectrice, j'espère que vous apprécierez ce moment passé à lire cette thèse et qu'elle avivera votre curiosité pour l'univers incroyable des symbioses.

Résumé

Les bactéries du sol appartenant au genre *Bradyrhizobium* ont la capacité de s'associer symbiotiquement avec une large diversité de légumineuses. Cette interaction conduit à la formation d'un nouvel organe racinaire, le nodule, au sein duquel les bactéries sont hébergées et fixent le diazote atmosphérique pour le bénéfice de la plante en échange de substrats carbonés issus de la photosynthèse. Pendant longtemps, il a été supposé que l'interaction symbiotique nécessitait que la plante hôte perçoive un signal synthétisé par la bactérie pour déclencher la nodulation : les facteurs Nod (FNs). Cependant, certaines légumineuses peuvent être nodulées par des souches de *Bradyrhizobium* à travers d'autres voies, qui n'impliquent pas la perception des FNs. C'est le cas de certaines espèces du genre *Aeschynomene* qui sont nodulées par des souches de *Bradyrhizobium* photosynthétiques dépourvues des gènes canoniques *nodABC* nécessaires à la synthèse des FNs. Plus récemment, un second processus indépendant des FNs a été identifié entre des souches de *Bradyrhizobium* non-photosynthétiques et des légumineuses tropicales, qui diffère du premier en raison du recrutement d'un système de sécrétion de type III (T3SS). Il a été montré que durant l'interaction entre *B. vignae* ORS3257 et *A. indica*, un nouvel effecteur de type III (T3E) nucléaire nommé ErnA est l'acteur principal déclenchant l'organogénèse nodulaire. Il fait partie d'une nouvelle famille de T3Es, appelée ET-Nod pour "Effector Triggering Nodulation". L'objectif principal de ce travail de thèse a été d'avancer dans la compréhension de ces processus symbiotiques indépendants des FNs. Pour cela, les propriétés symbiotiques sur *A. indica* d'une large diversité de *Bradyrhizobium* séquencés ont été analysées. Il a ainsi été montré que la symbiose T3SS-dépendante est répandue chez les *Bradyrhizobium* non-photosynthétiques. Par ailleurs, une analyse génomique comparative couplée à des approches de mutagénèse a permis d'identifier divers ET-Nods distincts d'ErnA, dont Sup3 codant pour une SUMO-protéase. En parallèle de ces travaux, nous avons montré que l'espèce modèle *A. evenia* lignée 21 est parfaitement adaptée pour conduire des études fonctionnelles sur cette symbiose T3SS-dépendante. De plus, les tests symbiotiques réalisés sur des mutants EMS d'*A. evenia* lignée 76 affectés dans les gènes requis pour la symbiose Nod-indépendante, T3SS-indépendante nous permettent d'avoir un premier aperçu de la voie de signalisation empruntée par le processus T3SS-dépendant, où les gènes symbiotiques requis varient selon les souches considérées et leur contenu en ET-Nods. Enfin, un dernier chapitre de cette thèse a consisté à analyser plus particulièrement le phylogroupe des *Bradyrhizobium* photosynthétiques dont plus de 200 nouvelles souches isolées de la rhizosphère du riz ont récemment été séquencées. Il est apparu que le caractère photosynthétique n'est finalement pas un trait universel de ce phylogroupe mais que l'ensemble de ces nouvelles bactéries testées sont capables de noduler *A. indica*. Ce dernier résultat suggère que les déterminants gouvernant la symbiose Nod-indépendante, T3SS-indépendante font partie du core-génome de ces bactéries et sont essentiels à leur survie.

L'ensemble des travaux de cette thèse apporte ainsi une nouvelle vision du rôle du T3SS dans la symbiose fixatrice d'azote où divers ET-Nods sont capables de court-circuiter les premières étapes de la signalisation symbiotique médiées par les FNs. L'émergence d'un nouveau modèle plante pour l'étude de cette symbiose atypique ouvre par ailleurs la voie à de nombreuses approches fonctionnelles qui permettront de déchiffrer les bases moléculaires par lesquelles les ET-Nods activent l'organogénèse nodulaire. Enfin, de nouvelles perspectives émergent également pour progresser dans la compréhension des mécanismes symbiotiques Nod-indépendant, T3SS-indépendant.

Mots clés : T3SS, effecteur, fixation symbiotique de l'azote, légumineuses, *Bradyrhizobium*.

Abstract

Soil bacteria belonging to the *Bradyrhizobium* genus have the ability to associate symbiotically with a wide variety of leguminous plants. This interaction leads to the formation of a new root organ, the nodule, in which the bacteria are housed and fix atmospheric nitrogen for the benefit of the plant in exchange for carbon substrates derived from photosynthesis. For a long time, it was assumed that the symbiotic interaction required the plant to perceive a signal synthesised by the bacteria in order to trigger nodulation: Nod factors (NFs). However, some legumes can be nodulated by *Bradyrhizobium* strains via other pathways that do not involve the perception of NFs. This is the case for certain species of the genus *Aeschynomene*, which are nodulated by photosynthetic *Bradyrhizobium* strains lacking the canonical *nodABC* genes required for the synthesis of NFs. More recently, a second Nod-independent process has been identified between non-photosynthetic *Bradyrhizobium* strains and tropical legumes, which differs from the first in that it involves the recruitment of a type III secretion system (T3SS). It has been shown that during the interaction between *B. vignae* ORS3257 and *A. indica*, a new type III nuclear effector (T3E) called ErnA is the main player triggering nodule organogenesis and is part of a new family of T3Es, called ET-Nod for "Effector Triggering Nodulation". This thesis mainly aimed at advancing our understanding of these symbiotic processes, which are independent of NFs. To this end, the symbiotic properties on *A. indica* of a wide diversity of sequenced *Bradyrhizobium* were analysed. We observed that T3SS-dependent symbiosis is widespread in non-photosynthetic *Bradyrhizobium* and a comparative genomic analysis coupled with mutagenesis approaches made it possible to identify various ET-Nods distinct from ErnA, including Sup3 encoding a SUMO-protease. Alongside this work, we were able to show that the model species *A. evenia* line 21 is perfectly suited for conducting functional studies on this T3SS-dependent symbiosis. In addition, symbiotic tests carried out on EMS mutants of *A. evenia* line 76 affected in genes required for Nod-independent, T3SS-independent symbiosis have given us an initial insight into the signalling pathway used by the T3SS-dependent process, where the symbiotic genes required vary depending on the strains considered and their ET-Nod content. Finally, in a final chapter of this thesis, we analysed more closely the photosynthetic *Bradyrhizobium* phylogroup, of which more than 200 new strains isolated from the rhizosphere of rice were recently sequenced. It emerged that the photosynthetic character is not a universal trait of this phylogroup, but that all the new bacteria tested are capable to nodulate *A. indica*. This latest result suggests that the determinants governing the Nod-independent, T3SS-independent symbiosis are part of the core-genome of these bacteria and are essential to their survival.

All the work in this thesis thus provides a new vision of the role of the T3SS in nitrogen-fixing symbiosis, where various ET-Nods are capable of by-passing the first stages of symbiotic signalling mediated by NFs. The emergence of a new plant model for the study of this atypical symbiosis also paves the way to numerous functional approaches that will allow us to decipher the molecular basis by which ET-Nods activate nodule organogenesis. Finally, new perspectives are also emerging to help better understand Nod-independent, T3SS-independent symbiotic mechanisms.

Key words: T3SS, effector, nitrogen-fixing symbiosis, legumes, *Bradyrhizobium*.

Table des matières

Introduction générale **1**

1. Généralités sur l'azote **1**

1.1. L'azote, un élément limitant pour la production végétale 2

1.1.1. Révolution verte et procédé Haber-Bosch 2

1.1.2. Contexte démographique 3

1.1.3. Association rhizobium-légumineuse : une alternative 4

1.2. La fixation biologique de l'azote 4

1.2.1. Le complexe nitrogénase 5

1.2.2. Fixation symbiotique 6

2. Généralités sur la symbiose rhizobium-légumineuse **8**

2.1. Le partenaire bactérien 8

2.1.1. Les rhizobiums 8

2.1.2. Un set de gènes important pour l'interaction symbiotique 8

2.1.3. Un double mode de vie : libre ou symbiotique 9

2.1.4. Les *Bradyrhizobium* : le genre bactérien le plus abondant sur Terre 10

2.2. Le partenaire plante 11

2.2.1. Les légumineuses 11

2.2.2. Intérêts agronomiques, écologiques et économiques 12

2.2.3. Les Papilionoideae : la plus large sous-famille des Fabacées 13

3. La symbiose nodulaire : de l'infection de la plante par la bactérie à la fixation d'azote **14**

3.1. Différents modes d'infection 14

3.1.1. Infection intracellulaire par cordon d'infection 14

3.1.2. Infection intercellulaire par crack-entry 15

3.2. Mise en place de l'organe symbiotique : le nodule 16

3.2.1. Division cellulaire et organogenèse nodulaire 16

3.2.2. Contrôle moléculaire du développement nodulaire 18

3.2.3. Internalisation des bactéries au sein du symbiosome et différenciation en bactéroïdes 20

3.3. Importance du dialogue moléculaire rhizobium-légumineuse 21

3.3.1. Les flavonoïdes 21

3.3.2. Les facteurs Nod 22

3.4. Cascade de signalisation 24

3.4.1. Un complexe de récepteur des FNs 24

3.4.2. Déclenchement et décryptage du signal calcique 26

3.4.3. Expression des premiers gènes de nodulation 27
3.4.4. Similarités avec la symbiose mycorhizienne : voie de signalisation symbiotique commune 28
3.5. Métabolisme du nodule 30
3.5.1. Fixation d'azote par la nitrogénase et taux d'oxygène 30
3.5.2. Auxotrophie symbiotique 31
3.5.3. Echanges nutritionnels entre la plante et la bactérie 31
4. L'immunité de la plante : un équilibre fragile entre infection et résistance 33
4.1. Généralités sur l'immunité de la plante : contexte d'une plante résistante 33
4.1.1. Immunité à la surface des cellules : première ligne de défense 33
4.1.2. Immunité intracellulaire : un second niveau de défense en réponse à l'infection 34
4.1.3. Coévolution des gènes de résistance et des effecteurs : le modèle en zig-zag 36
4.2. Spécificité d'interaction 36
4.2.1. Spectre d'hôte : les déterminants bactériens 36
4.2.2. Spectre d'hôte : les déterminants plantes 38
5. Existence de symbioses nodulaires atypiques 40
5.1. Le couple symbiotique *Bradyrhizobium*/*Aeschynomene* spp 40
5.1.1. Les *Aeschynomene* : généralités 40
5.1.2. Des légumineuses à nodulation de tige 41
5.1.3. Les *Bradyrhizobium* photosynthétiques 42
5.2. Le mécanisme Nod-indépendant, T3SS-indépendant entre les *Bradyrhizobium* photosynthétiques et les *Aeschynomene* spp. du groupe 3 44
5.2.1. Un processus symbiotique aussi efficient que la symbiose Nod-dépendante 45
5.2.2. La recherche du signal "non Nod" : approche de mutagénèse bactérienne 46
5.2.3. La recherche des déterminants symbiotiques du côté plante : le modèle *Aeschynomene evenia* 47
5.2.4. Un complexe de récepteur reconnaissant le signal "non Nod" ? 48
5.2.5. Divers scénarios évolutifs de la symbiose Nod-indépendante 49
5.3. Découverte d'un second processus symbiotique Nod-indépendant 50
5.3.1. Le processus Nod-indépendant, T3SS-dépendant chez le soja et les *Aeschynomene* 50
5.3.2. Identification de T3Es déclenchant l'organogenèse nodulaire 52
6. Revue : The versatile roles of type III secretion systems in rhizobium-legume symbioses. 54
7. Présentation du sujet de thèse 55
Chapitre 1 58
1. Contexte de l'étude 58
2. Article 1 : "Widespread *Bradyrhizobium* distribution of diverse Type III effectors that trigger legume nodulation in the absence of Nod factor" 59

Chapitre 2 60

1. Contexte de l'étude 60

2. Article 2 : "New insights into the T3SS-dependent symbiotic pathway between *Aeschynomene evenia* and *Bradyrhizobium* strains" 61

Chapitre 3 62

Caractérisation fonctionnelle des ET-Nods (projet ANR "ET-Nod") 62

1. Contexte de l'étude 62

2. Avancements dans la caractérisation fonctionnelle d'ErnA 63

2.1. Recherche des cibles précoces régulées transcriptionnellement par ErnA 63

2.2. Etablir l'interactome d'ErnA 69

2.3. Interaction des ET-Nods avec des partenaires SUMOylés 71

Chapitre 4 73

1. Contexte de l'étude 73

2. Article 3 : "Phylogenetic and population structure of the *nod*-free but nodulating *Bradyrhizobium* phylogroup" 74

3. Données complémentaires : de nouveaux T3SS chez les *Bradyrhizobium* dont le rôle reste à déterminer 75

Discussion et perspectives 78

1. La symbiose Nod-indépendante, T3SS-indépendante : un processus très efficace mais restreint au supergroupe des *Bradyrhizobium* "photosynthétiques" 78

1.1. Nouvelles connaissances sur le phylogroupe des *Bradyrhizobium* "photosynthétiques" 78

1.2. Nouvelles pistes pour l'identification du signal "non Nod" 79

2. La symbiose Nod-indépendante, T3SS-dépendante : un processus plus largement répandu chez les *Bradyrhizobium* mais dont le mutualisme reste à démontrer 81

3. Diversité des ET-Nods 83

3.1. Des effecteurs distincts les uns des autres mais qui partagent certaines similitudes 83

3.2. Vers l'identification de nouveaux ET-Nods et de nouveaux processus Nod-indépendants 84

3.3. Les SUMO-protéases : un vivier d'ET-Nods ? 85

4. Mode d'action des ET-Nods 86

4.1. Lien entre ET-Nods et partenaires SUMOylés 86

4.2. ErnA : un ET-Nod agissant sur des régulateurs centraux de la plante 87

4.3. Des cibles candidates pour ErnA à explorer 88

5. Décryptage de la voie de signalisation empruntée par la symbiose T3SS-dépendante 91

6. Futurs défis et perspectives 92

6.1. Améliorer les performances symbiotiques des rhizobiums *via* les ET-Nods dans un contexte Nod-dépendant 92

6.2. Vers un transfert de la capacité symbiotique aux céréales ? 93

6.3. Les ET-Nods : un moyen de faciliter la nodulation chez les non-légumineuses ? 94
Références bibliographiques 96
Annexes 121
Annexe 1 : Articles complémentaires 121
Article : "Phylogenetic distribution and evolutionary dynamics of nod and T3SS genes in the genus *Bradyrhizobium*" 121
Article : "Identification of type III effectors modulating the symbiotic properties of *Bradyrhizobium vignae* strain ORS3257 with various *Vigna* species" 122
Article : "The Type III Effectome of the Symbiotic *Bradyrhizobium vignae* Strain ORS3257" 123
Annexe 2 : Matériels et méthodes complémentaires du chapitre 3 124
1.1. Matériels biologiques et conditions de culture 124
1.2. Constructions des plasmides et complémentation 124
1.2.1. Ajout des tags et complémentation du mutant *ORS3257ΔernA* 124
1.2.2. Création du système inductible pDEX-ernA 124
1.2.3. Obtention du mutant dans le site catalytique du domaine SUMO-protéase de Sup3 125
1.3. Analyse RNA-seq 125
1.4. Autres matériels et méthodes 126

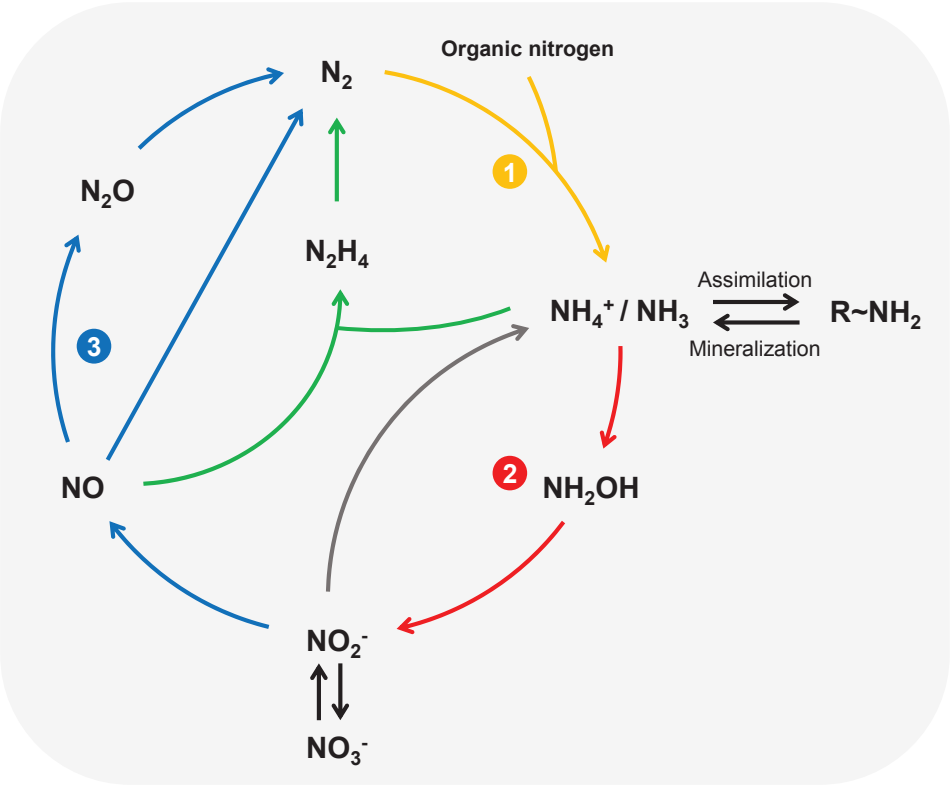


Figure 1. Principales étapes du cycle de l'azote (adapté de Stein & Klotz, 2016).
L'ensemble des réactions qui constituent le cycle de l'azote sont représentées par différentes flèches de couleurs. Les trois processus majeurs sont (1) la fixation de l'azote (jaune), (2) la nitrification (rouge) et (3) la dénitrification (bleu). A ces étapes viennent s'ajouter une étape d'ammonification de l'azote organique qui a lieu en parallèle de la fixation de l'azote (jaune), une étape de réduction dissimilaire des nitrites (NO_2^-) en ammonium (NH_4^+) (gris), une étape d'oxydation anaérobique de l'ammonium appelée anammox (vert) et celle de l'interconversion nitrite-nitrate (noir). Ce sont l'ammonium (NH_4^+) et le nitrate (NO_3^-) qui vont pouvoir être assimilés par les plantes.

Introduction générale

1. Généralités sur l'azote

L'azote[1] est un élément majeur de notre atmosphère et se trouve être un des composants essentiels du vivant avec le carbone, l'oxygène et l'hydrogène. En effet, l'azote ou diazote (N_2) est le constituant majoritaire de l'atmosphère terrestre (≈78%). Il est indispensable au vivant puisqu'il est présent dans l'ADN et l'ARN à travers les acides nucléiques formées en partie de bases azotées et se trouve donc être un élément universel à la vie. Sous forme d'acides aminés (groupement amine -NH_2) il entre dans la composition des protéines (≈15%) qui constituent les principaux composants cellulaires et qui sont aussi essentielles à la nutrition des organismes vivants dans leur ensemble.

Paradoxalement, le N_2 abondant de l'atmosphère n'est pas biodisponible pour la grande majorité des êtres vivants et se trouve être assimilable que par certains procaryotes. C'est à travers le cycle biogéochimique de l'azote que le N_2 va être transformé en diverses formes réactives de l'azote et va ainsi devenir assimilable par l'ensemble des organismes vivants sur Terre. Trois processus majeurs constituent le cycle de l'azote : i) une étape de fixation du N_2 qui peut être symbiotique ou asymbiotique, ii) la nitrification conduisant à la production de nitrate et iii) la dénitrification qui rendra à l'atmosphère l'azote sous sa forme moléculaire N_2 mais aussi émettra une faible quantité de protoxyde d'azote ou oxyde nitreux N_2O.

A noter que le N_2O est un important gaz à effet de serre et bien que sa production naturelle soit faible, l'impact de l'homme sur ces émissions font de l'oxyde nitreux l'un des gaz à effet de serre majeur à réduire au 21ème siècle et dans le futur (Stein & Klotz, 2016). Son effet sur la destruction de la couche d'ozone est considéré comme 300 fois plus puissant que celui du dioxyde de carbone (CO_2) et joue donc un rôle majeur sur le réchauffement climatique. De récentes études ont permis d'améliorer notre compréhension du cycle de l'azote qui comprend aujourd'hui les trois étapes originalement décrites auxquelles viennent s'ajouter l'ammonification qui intervient à l'étape de fixation du N_2, l'oxydation anaérobique de l'ammonium appelée aussi anammox, la réduction du nitrite en ammonium et l'interconversion nitrite-nitrate (Stein & Klotz, 2016) (figure 1). Fait remarquable, l'ensemble de ces étapes fait intervenir des cohortes de microorganismes (bactéries, archées, champignons) soulignant leur importance dans le cycle de l'azote (Hayatsu et al., 2008). Ceux impliqués dans la fixation directe de l'azote sont appelés diazotrophes et sont strictement limités aux procaryotes.

[1] Du composé a- (privatif) et du radical grec ζωτ- "vivant", signifiant donc "privé de vie".

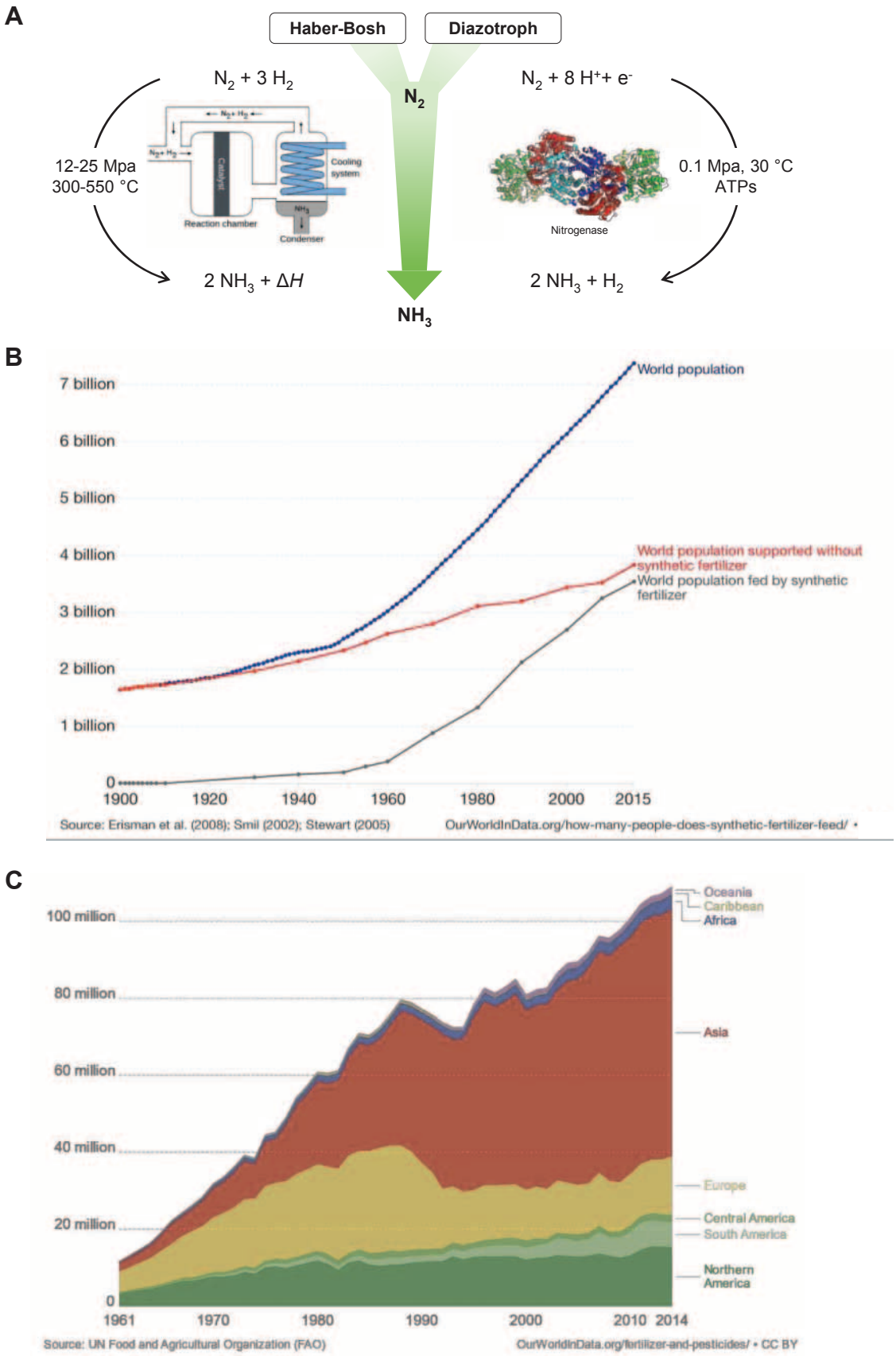


Figure 2. Révolution de l'agriculture par le procédé Haber-Bosch et contexte démographique.
(**A**) Réactions chimiques permettant la fixation de l'azote atmosphérique. Dans le procédé Haber-Bosh, elle nécessite une pression entre 12 et 25 Mpa et une température variant de 300 à 550 °C. En comparaison, les organismes diazotrophes ont besoin uniquement d'une pression de 1 bar (0,1 Mpa) et d'une température de 30 °C, le tout catalysé par de l'ATP comme source d'énergie (Rapson et al., 2020). (**B**) Estimation de la population mondiale jusqu'en 2015, avec ou sans engrais azotés issus du procédé Haber-Bosch pour la production alimentaire. La meilleure des estimations prédit que seulement la moitié de la population mondiale pourrait être nourri en absence de l'utilisation d'engrais azotés. (**C**) Consommation des engrais azotés issus du procédé Haber-Bosch, en tonnes par année dans le monde entre 1961 et 2014.

1.1. L'azote, un élément limitant pour la production végétale

L'azote fait partie des éléments minéraux essentiels au bon développement des plantes avec notamment le phosphore et le potassium. La nutrition minérale de la plante passe par l'absorption de ces éléments par le système racinaire qui permet d'explorer le sol à la recherche de nutriments. Malheureusement, leurs formes assimilables par la plante sont souvent rares et difficilement biodisponibles. Dans le cas de l'azote, la plante n'est capable d'assimiler que les formes ioniques que sont le nitrate NO_3^- et l'ammonium NH_4^+ qui sont présentes qu'en très faibles quantités dans les sols. L'azote représente donc un élément limitant pour la croissance des plantes et par conséquent pour la production végétale.

1.1.1. Révolution verte et procédé Haber-Bosch

Le terme de "révolution verte" désigne une période faste de la production agricole suite aux progrès scientifiques réalisés notamment dans le domaine de la chimie, de la mécanisation et dans la mise au point par les semenciers de nouvelles variétés à haut rendement. Cette période fait aussi référence à une politique de transformation des pratiques agricoles dans les pays en développement au cours de la période 1960-1990. Cette forte augmentation des rendements agricoles a été possible grâce à l'utilisation massive d'engrais azotés issus du procédé Haber-Bosch qui a révolutionné l'agriculture en permettant de produire de l'ammoniac NH_3 à partir de l'hydrogénation du N_2 (figure 2A). Initialement mis en place à la veille de la première guerre mondiale pour la production de la poudre à canon, ce procédé n'a été que pleinement utilisé dans les années 1950. En plus d'une augmentation spectaculaire de la productivité agricole, la révolution verte a eu pour conséquence d'éviter des famines dans les pays en cours de développement et les moins avancés et d'accroître considérablement la population mondiale. L'agriculture d'aujourd'hui a hérité de ce que nous a apporté la révolution verte et a donné naissance au système agricole industrialisé que nous connaissons. C'est environ la moitié de la population mondiale qui est nourrit grâce aux engrais produits par le procédé Haber-Bosch (figure 2B). Cependant, ce procédé est très couteux en énergie : il nécessite une forte température et une pression élevée (Erisman et al., 2008). En comparaison, les organismes diazotrophes catalysent la même réaction ($N_2 \rightarrow NH_3$) à faible pression et température, avec comme source d'énergie de l'ATP[2] (figure 2A ; Rapson et al., 2020). La population en continuelle croissance a poussé à une production et une utilisation des engrais azotés chimiques de plus en plus forte (figure 2C). Cela conduit à un dérèglement du cycle de l'azote avec une pollution en NO_3^- considérable car seule une portion de cet azote est réellement consommée par les plantes. Très soluble dans l'eau, le sur-

[2] Adénosine-triphosphate.

plus de NO_3^- est responsable de la contamination des grands réservoirs d'eau souterrains du globe et peut avoir des conséquences sur la santé humaine (Abascal et al., 2022). Le lessivage de ces engrais sous forme de NO_3^- dans les eaux fluviales est aussi responsable de l'eutrophisation des eaux par les algues et les microorganismes, qui réduisent l'accès à l'oxygène aux autres organismes vivants dans ces écosystèmes : toute autre forme de vie disparait donnant naissance à ce qu'on appelle des zones côtières mortes (Erisman et al., 2013). Une pollution au N_2O est aussi observée suite à l'utilisation d'engrais qui a un fort impact sur le réchauffement climatique et conduit à la destruction de la couche d'ozone (Stein & Klotz, 2016). L'utilisation des engrais produit par le procédé Haber-Bosch altère donc considérablement le cycle de l'azote global qui apparaît être un enjeu sociétal majeur car il peut se trouver perturber par l'activité humaine et avoir un fort impact sur l'agriculture, le réchauffement climatique mais aussi sur la santé humaine (Townsend et al., 2003; Galloway et al., 2008; Schlesinger, 2009; Johnson et al., 2010; Tian et al., 2018;).

Comprendre les besoins nutritifs d'une plante

Les plantes sont des organismes photoautotrophes qui utilisent la lumière comme source d'énergie et le CO_2 comme source de carbone. Le métabolisme carboné est donc d'une importance centrale chez les végétaux. Les autres macroéléments (hydrogène, oxygène, azote, potassium, calcium, magnésium, phosphore et soufre) et microéléments (chlore, fer, bore, manganèse, zinc, cuivre, nickel, silicium et molybdène) nécessaires à sa nutrition minérale sont puisés directement dans le sol grâce à son système racinaire ou absorbés de manière gazeuse à travers les stomates situés sur les feuilles. En raison de leur mode de vie immobile, les plantes ont accès à des ressources nutritives diluées et fluctuantes et ont pour cela développé de nombreuses spécificités pour leur adaptation (grande surface d'échange foliaire et racinaire, développement plastique, systèmes de transport à forte affinité, interaction positives avec des microorganismes etc.). En agriculture, pour subvenir aux besoins nutritifs des plantes et assurer un bon rendement agricole, les mélanges d'engrais chimique NPK sont les plus largement utilisés, auxquels peuvent s'ajouter d'autres engrais organiques pour apporter les autres éléments nécessaires à leur croissance.

1.1.2. Contexte démographique

En 2022, la population mondiale a atteint les 8 milliards d'habitants et devrait culminer autour des 10 milliards en 2080 (www.un.org). L'Inde et la Chine étant actuellement les pays les plus peuplés du monde, c'est cependant en Afrique subsaharienne que l'augmentation de la population d'ici 2050 devrait être la plus forte. L'augmentation de la production alimentaire est donc l'un des principaux défis de notre siècle pour parvenir à une sécurité alimentaire mondiale. C'est pourquoi de nombreuses recherches et études sont menées dans le domaine de l'agriculture pour améliorer la production végétale. Dans ce contexte, la compétition pour de nouvelles terres agricoles conduit à une concurrence croissante pour l'eau et l'énergie qui affecte à la fois les capacités de production mais aussi l'impact environnemental qui en découle (Godfray et al., 2010; Frona et al., 2019). Pour garan-

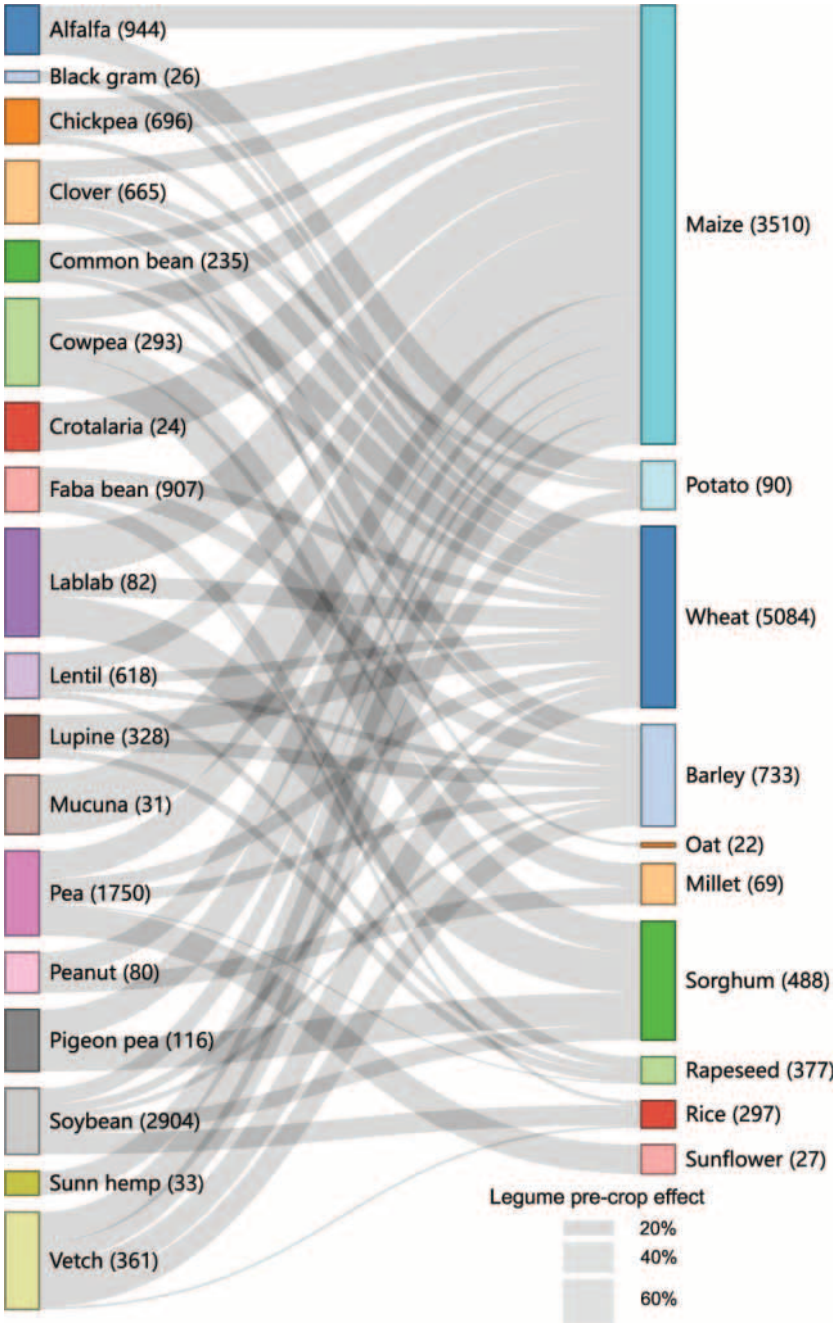


Figure 3. Effet de la rotation culturale avec des légumineuses sur le rendement des céréales (Zhao et al., 2022).
La largeur des flux entre chaque légumineuse-céréale représente l'amplitude de l'effet de la légumineuse avant la culture (à gauche), sur le rendement de la culture de la céréale (à droite). Les séquences de cultures sont indiquées lorsque le nombre d'observations est supérieur à 20 (indiqué entre parenthèse). Une grande diversité de légumineuses pré-cultivées (ici 18) est capable d'améliorer le rendement lors de la culture principale. Les céréales de fort intérêt agronomique que sont le maïs, le blé et le riz voient leur rendement amélioré par la pré-culture d'au moins trois espèces de légumineuses différentes, dont au moins une peut être aussi utilisée pour la production alimentaire (par exemple le soja qui augmente le rendement du riz de plus de 20%).

tir une sécurité alimentaire durable et équitable, il est donc nécessaire de repenser à notre manière de production agricole pour qu'elle puisse répondre à l'augmentation de la population tout en réduisant notre impact sur l'environnement. Une nouvelle forme de révolution verte cette fois-ci moins polluante, est donc essentielle pour espérer subvenir aux besoins alimentaires de notre population avec une attention particulière au continent Africain.

1.1.3. Association rhizobium-légumineuse : une alternative

Les chapitres précédents montrent qu'il est nécessaire de repenser à notre façon de produire pour répondre aux défis actuels (X. Zhang et al., 2015). La pratique de rotation culturale déjà utilisée au Moyen-Âge est un concept qui est aujourd'hui réintroduit dans les pratiques modernes d'agriculture durable pour ses nombreux avantages en comparaisons des systèmes de monocultures qui entraînent des effets négatifs tels que l'épuisement de certains éléments nutritifs dans les sols, la baisse de biodiversité mais aussi une réduction des rendements sur le long terme (Bourke et al., 2021). Parmi les rotations culturales pratiquées, celles qui introduisent les légumineuses sont les plus utilisées pour répondre au besoin azoté des plantes. En effet, leur capacité à s'associer avec des bactéries du sol, appelées rhizobiums, permet de rendre biodisponible pour la plante l'azote atmosphérique N_2 qui est converti en NH_4^+. Il a été ainsi montré que l'utilisation en agriculture de l'association symbiotique rhizobium-légumineuse permet d'augmenter les rendements d'environ 20% (figure 3), en particulier lorsque peu d'engrais azotés est utilisé (Zhao et al., 2022). Les avantages les plus importants de l'apport des légumineuses aux systèmes agricoles ont été observés en Afrique là ou l'utilisation d'intrants est la plus faible. Ainsi diversifier les cultures par les légumineuses est une des alternatives qui pourra dans le futur réduire les effets néfastes de l'utilisation massive d'engrais azotés, ce qui souligne l'importance de mieux comprendre comment l'interaction rhizobium-légumineuse fonctionne. Enfin, un des défis majeurs du 21ème siècle est de transférer cette capacité symbiotique, aujourd'hui restreinte aux légumineuses, aux graminées qui sont largement consommées dans le monde. Cette fixation d'azote aux céréales ne pourra être réussie qu'en déchiffrant l'ensemble des mécanismes sous-jacents de cette interaction et en les activant chez les céréales. Dans ce sens, le consortium ENSA[3] réuni de nombreux laboratoires et équipes de recherche pour l'ingénierie de la symbiose fixatrice d'azote par les rhizobiums (www.ensa.ac.uk).

1.2. La fixation biologique de l'azote

La fixation de l'azote est un processus clé du cycle de l'azote. A l'échelle mondiale, la moitié de l'azote fixé est apportée par les activités humaines à travers par exemple la production d'engrais

[3] Enabling Nutrient Symbioses in Agriculture.

Figure 4. Exemple d'organismes diazotrophes capables de fixer biologiquement l'azote à l'état libre.
(A) *Methanothermobacter thermautotrophicus* (archée). (B) *Pseudomonas stutzeri* (bactérie). (C) *Cylindrospermum stagnale* (cyanobactérie).

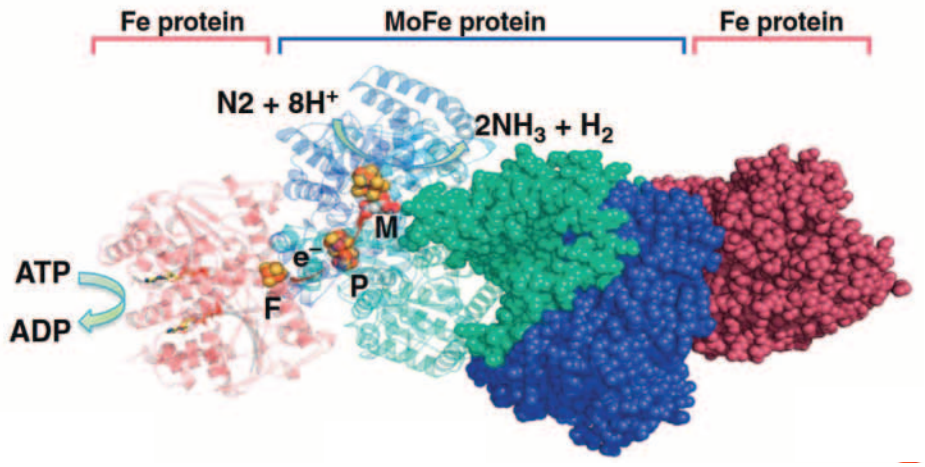


Figure 5. Représentation de la structure du complexe enzymatique de la nitrogénase (Oldroyd & Dixon, 2014).
Les deux composants de la nitrogénase sont représentés en bleu (MoFe protein) et en rouge (Fe protein). Durant un cycle catalytique, les deux composants s'associent puis se dissocient impliquant un transfert d'électrons et l'hydrolyse d'ATP. Les trois groupes métalliques F [4Fe-4S], P [8Fe-7S] et M (FeMo-co) sont indiqués dans la partie gauche du schéma. L'hydrolyse de l'ATP par la protéine Fe, le transfert d'électron associés jusqu'au site catalytique et la réaction enzymatique sont également représentés dans la partie gauche de la structure.

azotés, l'agriculture et l'élevage ou la combustion de biomasse (Fowler et al., 2013). C'est donc l'autre moitié qui est produit de manière biologique et non par l'homme. Cette source d'azote provient majoritairement des écosystèmes marins (⅔) et terrestres (⅓) avec un très faible apport par la foudre (≈2,4%). En absence de l'influence humaine, la fixation biologique de l'azote et la production de NO par la foudre seraient les seuls apports d'azote à l'environnement.

1.2.1. Le complexe nitrogénase

La capacité des procaryotes à fixer l'azote atmosphérique est restreinte à un petit groupe de bactéries, d'archées et de cyanobactéries aussi appelées "algues bleues" (figure 4). Cette fixation biologique de l'azote peut se faire à l'état libre ou en vie symbiotique selon les micro-organismes. Ils sont appelés diazotrophes car capables de réduire le N_2 en NH_3 grâce au complexe enzymatique de la nitrogénase. La réduction biologique de l'azote fait partie d'un des processus les plus couteux en énergie parmi l'ensemble des réactions enzymatiques existantes : 16 molécules d'ATP sont nécessaires par molécule de N_2 réduit. Le complexe enzymatique de la nitrogénase (≈220 kDa[4]) est un système à deux composants (figure 5). Il est constitué de la protéine MoFe (appelée dinitrogénase ou composant 1) et de la protéine Fe permettant le transfert d'électrons (appelée dinitrogénase réductase ou composant 2) (Oldroyd & Dixon, 2014).

Les protéines de structure de la nitrogénase sont codées respectivement par le gène *nifH* pour le composant 1 et les gènes *nifD* et *nifK* pour le composant 2. Ces trois gènes sont retrouvés chez tous les diazotrophes mais ne sont pas suffisants pour obtenir un complexe actif de la nitrogénase. Pour le bon fonctionnement du site catalytique, une source réductrice capable de donner des électrons et du Magnésium (Mg) ATP sont nécessaires. En effet, les protéines MoFe et Fe s'associent et se dissocient durant un cycle catalytique qui implique le transfert d'un électron et l'hydrolyse du MgATP. La protéine MoFe s'avère être elle-même constituée d'un cofacteur fer-molybdène (FeMo-co ou groupe M) qui fournit le site actif pour la fixation et la réduction du diazote, d'un groupe F [4Fe-4S] et d'un groupe P [8Fe-7S] impliqué dans le transfert d'électrons de la protéine Fe au FeMo-co[5] (figure 5). Par ailleurs en raison des propriétés oxydantes de l'oxygène, la nitrogénase voit son activité inhibée lorsque la pression partielle en oxygène devient trop élevée par la dégradation du Fe-S contenu dans le composant 2 alors incapable de transférer les électrons au composant 1. D'autres types de nitrogénases ont été découverts où le Mo du co-facteur est remplacé par du vanadium (V-nitrogénase) ou du fer (Fe-nitrogénase) chez certaines espèce d'*Azotobacter* (Eady, 1996). C'est cependant la Mo-nitrogénase qui est retrouvée chez les rhizobiums associés aux légumineuses.

[4] Kilodalton : unité de masse atomique.
[5] Les groupes M, F et P forment 3 groupes métalliques (en anglais "metalloclusters ").

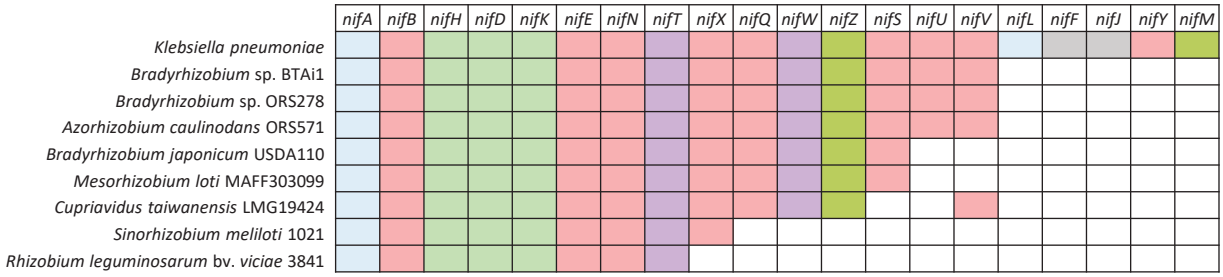


Figure 6. Présence ou absence des gènes *nif* chez la bactérie diazotrophe libre *Klebsiella pneumoniae* et chez certains rhizobiums (adapté de Masson-Boivin et al, 2009).
Les gènes présents sont colorés tandis que les gènes absents sont en blancs. Leur relative fonction dans le complexe enzymatique de la nitrogénase est indiquée selon le code couleur suivant : gène de régulation (bleu), gènes de structure de la nitrogénase (vert clair), gènes de maturation de la nitrogénase (vert foncé), gènes de biosynthèse du FeMo-co (rose), gènes de fonction inconnue (violet) et gènes impliqué dans le don d'électrons (gris). Chez les rhizobiums, FixU est remplacé par NifT.

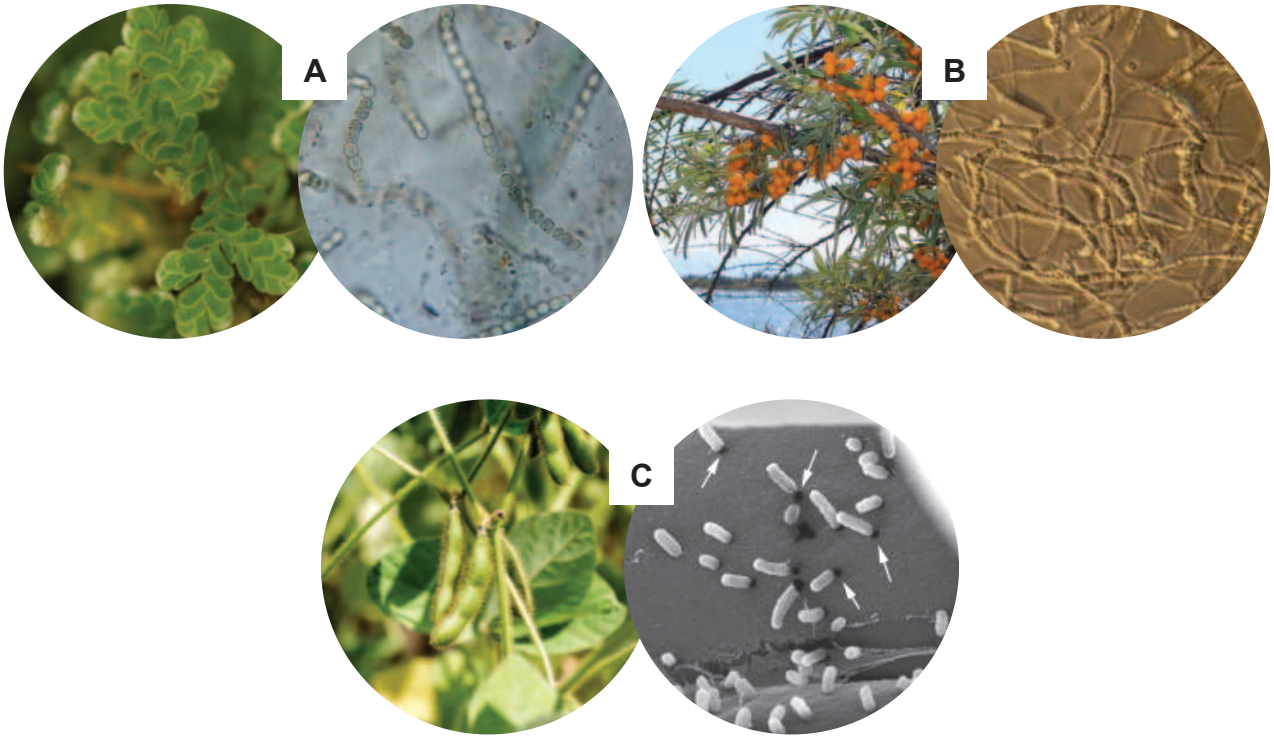


Figure 7. Illustration d'interactions plante/microorganisme capables de fixer symbiotiquement l'azote.
(**A**) Couple symbiotique de la fougère aquatique *Azolla* (à gauche) avec la cyanobactérie *Anabaena* (à droite) (source : https://digital.library.wisc.edu). (**B**) Couple symbiotique d'une plante actinorhizienne, l'argousier (à gauche) avec la bactérie filamenteuse *Frankia* (à droite) (source : Wikipedia). (**C**) Couple symbiotique du soja *Glycine max* (à gauche) avec un rhizobium (à droite) (source : https://www.sciencedirect.com)

L'activité du complexe enzymatique nécessite d'autres gènes *nif* jouant un rôle dans la biosynthèse et l'assemblage du FeMo-co ainsi que dans le transport des électrons et la régulation de la synthèse du complexe. L'analyse des génomes de certains diazotrophes a montré une grande variation dans le contenu en gènes *nif* (Masson-Boivin et al., 2009). De manière contre-intuitive, ce sont les rhizobiums qui montrent la plus grande réduction de ces gènes (figure 6), alors que ce sont eux qui jouent un rôle prépondérant dans la fixation biologique de l'azote (cf. 1.2.2). Par exemple, l'absence du gène *nifV* chez la grande majorité des rhizobiums empêche ces bactéries de fixer librement l'azote. En effet, *nifV* code pour une homocitrate synthase impliquée dans la production d'homocitrate qui rentre dans la composition du co-facteur FeMo-co essentiel pour le fonctionnement de la nitrogénase. Pour pallier à ce manque au cours de l'interaction symbiotique, c'est la plante, *via* le gène *fen1* codant pour une autre homocitrate synthase, qui fournit l'homocitrate à la bactérie (Hakoyama et al., 2009). De manière intéressante, certains rhizobiums du genre *Bradyrhizobium* et *Azorhizobium caulinodans* possèdent quant à eux le gène *nifV* et sont capables de fixer librement l'azote. A noter qu'au cours de la symbiose entre *Aeschynomene* et *Bradyrhizobium* photosynthétique (ORS285), il est observé une dépendance variable de *nifV* pour la fixation symbiotique de l'azote suivant l'espèce hôte (Nouwen et al., 2017). L'homocitrate est donc une molécule essentielle pour la bonne fixation symbiotique de l'azote et souligne la dépendance des rhizobiums vis-à-vis de leur plante hôte.

1.2.2. Fixation symbiotique

La fixation biologique de l'azote dans les terres agricoles produit une quantité considérable d'azote dans les sols, la majorité apportée par les associations symbiotiques entre les légumineuses et les rhizobiums mais aussi entre la cyanobactérie *Anabaena azollae* et des fougères aquatiques du genre *Azolla* retrouvées souvent dans les rizières (figure 7). L'autre partie de l'azote fixé sur les terres, cultivées ou non, est produite par des bactéries endophytes de plantes ou part des bactéries capables de fixer l'azote à l'état libre (cyanobactéries, bactéries hétérotrophes et phototrophes) (Ladha et al., 2022). Mais d'autres associations symbiotiques en plus de celles citées précédemment interviennent dans la fixation biologique de l'azote. Ce sont les associations entre les plantes actinorhiziennes et les bactéries du genre *Frankia* (figure 7) et celles entre les plantes de la famille des Cycadales avec des cyanobactéries du genre *Anabaena* (Herridge et al., 2008). Des associations entre champignons et diazotrophes ont également été reportées comme fixatrices de N_2 (Jayasinghearachchi & Seneviratne, 2004; Sapountzis et al., 2016). Cette interaction symbiotique plante/microorganisme peut être plus au moins intime entre les deux partenaires concernés ; allant de l'épiphytisme à l'endophytisme (Mus et al., 2016). Les microorganismes épiphytes se définissent par une colonisation de la plante en surface tandis que les microorganismes endophytes sont capables de pénétrer les

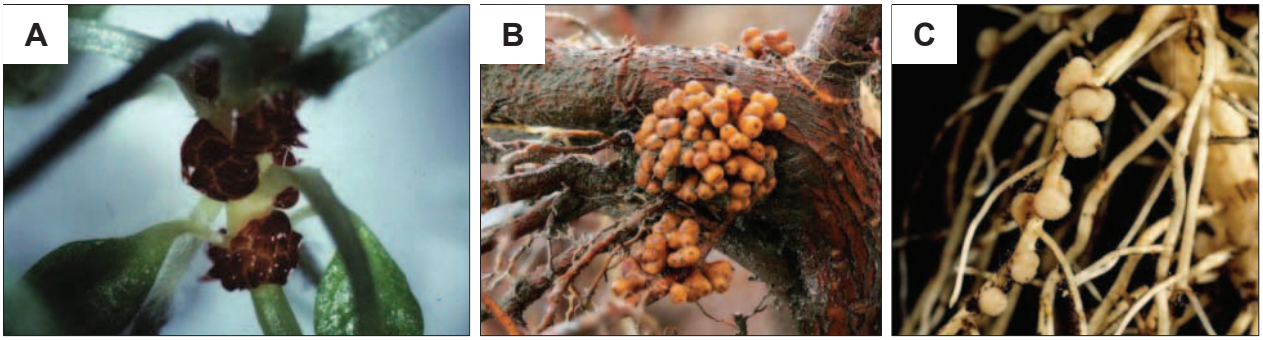


Figure 8. Illustration des trois endosymbioses fixatrices d'azote existantes chez les angiospermes.
(**A**) Association entre les cyanobactéries *Nostoc* sp. avec des espèces du genre *Gunnera* (Adams et al., 2013). On observe le développement de glandes rouges au niveau de la tige et à la base de chaque pétiole de la jeune plantule. Les cyanobactéries y sont hébergées intracellulairement. (**B**) Association entre des bactéries du genre *Frankia* et une plante actinorhizienne : l'aulne (Martin et al., 2017). Les bactéries sont hébergées à l'intérieur des nodules formés sur les racines de l'arbre. (**C**) Association entre des rhizobiums et la fève (*Vicia faba*) (Martin et al., 2017). Tout comme les *Frankia*, les rhizobiums sont hébergés au sein de nodules qui se développent généralement sur les racines de la plante.

tissus de la plante de manière intercellulaire ou intracellulaire (Hardoim et al., 2015). On parle de bactérie PGPR, pour Plant Growth Promoting (Rhizo)bacteria, lorsque la colonisation ne nécessite pas une pénétration des tissus (Ahemad, 2014) ; le terme PGPR incluant un ensemble de bactéries de la rhizosphère capables de coloniser l'environnement racinaire. On peut citer par exemple les PGPR appartenant au genre *Azospirillum* qui sont capables d'améliorer la croissance de plantes d'intérêts agronomiques telles que le blé, le maïs et le riz (Cassan & Diaz-Zorita, 2016). Ce sont cependant les bactéries endophytes pénétrant les tissus de la plante qui sont les plus efficientes pour la fixation d'azote et qui conduisent à la formation d'un nouvel organe chez la plante qui abrite le symbiote. On parle alors d'endosymbiose (Geurts et al., 2016). Chez les angiospermes, on peut distinguer trois types d'endosymbioses fixatrices d'azote (figure 8) : i) les cyanobactéries du genre *Nostoc* avec des espèces du genre *Gunnera* (Adams et al., 2013) ; ii) les bactéries actinomycètes du genre *Frankia* (Gram-positives) avec des plantes actinorhizienne (essentiellement des arbres ou arbustes) et iii) des α- et β-protéobactéries (Gram-négatives) communément appelées rhizobiums avec des légumineuses (famille des Fabacées) ou avec des plantes du genre *Parasponia* (famille des Cannabacées). A noter que *Parasponia* est la seule plante non-légumineuse s'associant avec des rhizobiums (Santi et al., 2013).

Qu'est-ce qu'une symbiose ?

Introduit en 1879 par Anton de Bary, pionnier dans l'étude des champignons et des algues, le terme de symbiose qui vient du grec συμβίωσις, sumbiōsis "vie en commun", a pour définition originale "une association vivante d'espèces différentes". Cette définition n'incluait donc pas l'idée qu'il puisse y avoir des bénéfices pour l'un ou pour l'ensemble des partenaires. Aujourd'hui couramment utilisé, le terme de symbiose englobe tout type d'associations interspécifiques, qu'elles soient profitables ou non aux organismes associés aussi appelés symbiotes. Ces associations peuvent appartenir à cinq catégories différentes (Veiga, 2016) :

- Le mutualisme : une interaction où les deux symbiotes tirent avantages (+/+).
- Le commensalisme : une interaction bénéfique à un seul des symbiotes mais qui n'affecte pas l'autre (+/0).
- L'amensalisme : une interaction neutre où aucun effet n'est observé (0/0).
- Le parasitisme : une interaction qui avantage un des symbiotes au détriment de l'autre (+/-).
- La synnécrose : une interaction conduisant à la mort des deux symbiotes (-/-).

Néanmoins lorsqu'on désigne une symbiose, on l'associe le plus généralement à une interaction qui bénéficie à au moins l'un des partenaires sans affecter l'autre. Enfin selon la biologiste Lynn Margulis, l'endosymbiose (c'est-à-dire lorsque le symbiote vit à l'intérieur de son hôte) serait un facteur clé de l'évolution des espèces. Les eucaryotes seraient le résultat évolutif d'associations symbiotiques avec divers procaryotes ou autres organismes unicellulaires ; mettant en avant la notion de coopération et de dépendance entres les organismes vivants. Dans ce sens, les organites tels que les chloroplastes et les mitochondries sont devenues des symbioses obligatoires et ont été acquis au cours de l'évolution par endosymbiose avec des bactéries. Ces organites possèdent leur propre génome et sont nécessaires au métabolisme cellulaire. Un parallèle entre ces symbioses devenues obligatoires et celle de la symbiose fixatrice d'azote peut être fait au regard de l'existence d'un organite temporaire mis en place lors de l'interaction pour fixer le diazote : le symbiosome (Coba de la Peña et al., 2018).

Family	Genus	Species of rhizobia within the genus
Rhizobiaceae	*Rhizobium*	133 species
	Ensifer/*Sinorhizobium*	≈24 species
	Allorhizobium	9 species
	Pararhizobium	*P. giardinii* and *P. herbae*
	Shinella	*S. kummerowiae*
	Neorhizobium	*N. alkalisoli*, *N. galegae* and *N. huautlense*
Hyphomicrobiaceae	*Devosia*	*D. neptuniae*
Bradyrhizobiaceae	*Bradyrhizobium*	60 species
	Blastobacter	*B. denitrificans*
Phyllobacteriaceae	*Mesorhizobium*	59 species
	Aminobacter	*A. anthyllidis*
	Phyllobacterium	8 species
Methylobacteriaceae	*Methylobacterium*	*M. nodulans*
	Microvirga	*M. lupini*, *M. lotonononidis*, *M. zambiensis* and *M. vignae*
Brucellaceae	*Ochrobactrum*	*O. lupini* (synonym of *O. anthropic*), *O. cytisi*
Burkholderiaceae	*Paraburkholderia*	*Paraburkholderia* 69 species
	Cupriavidus	*C. taiwanenses*
	Trinickia	*T. symbiotica*
	Ralstonia	*R. taiwanensis* (Synonymous: *Cupriavidus taiwanensis*; *Wautersia taiwanensis*)
Xanthobacteraceae	*Azorhizobium*	*A. caulinodans*, *A. doebereinerae* and *A. oxalatiphilum*

Tableau 1. Classification et nombre d'espèces des bactéries identifiées comme rhizobium (adapté de Moura et al., 2020 et Santos et al., 2018).
L'ensemble des rhizobiums sont distribués dans le genre des α-protéobactéries (Rhizobiaceae, Phyllobacteriaceae, Brucellaceae, Methylobacteriaceae, Bradyrhizobiaceae, Xanthobacteraceae et Hyphomicrobiaceae) et des β-protéobactéries (uniquement dans la famille des Burkholderiaceae). Tous les rhizobiums du genre *Burkholderia* (32 espèces) on été reclassés en *Paraburkholderia*.

2. Généralités sur la symbiose rhizobium-légumineuse

Parmi les trois endosymbioses bactériennes citées précédemment, celles impliquant les rhizobiums et les *Frankia* conduisent à la formation d'un nouvel organe, le plus souvent racinaire, appelé nodule rhizobien pour les rhizobiums ou nodule actinorhizien lorsqu'il s'agit des *Frankia*. On parle alors d'endosymbiose nodulaire, où les symbiotes sont hébergés dans le nodule et fixent l'azote atmosphérique pour le bénéfice de la plante. En contrepartie, la plante fournit aux bactéries une niche écologique et un apport de substrats carbonés issus de la photosynthèse. Dans les paragraphes qui vont suivre, nous allons nous intéresser plus particulièrement à la symbiose rhizobium-légumineuse.

2.1. Le partenaire bactérien

2.1.1. Les rhizobiums

Le terme général de "rhizobium" ne représente pas un taxon unique mais recouvre l'ensemble des bactéries capables d'induire des nodules (racinaire ou caulinaire) chez leur plante hôte : les légumineuses. Le terme rhizobium désigne ainsi un groupe polyphylétique de plusieurs lignées de bactéries Gram-négatives au sein des protéobactéries ayant une fonction similaire (De Lajudie & Young, 2017). Aujourd'hui on compte 8 familles de rhizobiums et 20 genres différents englobant plus de 400 espèces (tableau 1). La plupart des symbiotes sont répartis dans la classe des α-protéobactéries. Néanmoins, il existe aussi d'autres symbiotes mais moins nombreux dans la classe des β-protéobactéries (Masson-Boivin et al., 2009; De Lajudie et al., 2019; Moura et al., 2020).
Cette diversité est en perpétuelle évolution du fait de la découverte de nouvelles souches et ces chiffres vont certainement continuer d'augmenter (De Lajudie et al., 2019; Moura et al., 2020). En effet, l'ère des nouveaux outils de séquençage massif des génomes a permis et permet d'identifier de nouveaux rhizobiums notamment par l'identification des gènes *nod* et *nif* dans leur génome. Ainsi, chaque année se tient un comité internationale sur la taxonomie des rhizobiums (ICSP) afin de mettre à jour les données nouvelles apportées et définir de nouveaux sous genres et espèces (Mousavi & Young, 2022).

2.1.2. Un set de gènes important pour l'interaction symbiotique

Génétiquement parlant, les rhizobiums possèdent un groupe de gènes spécifiques qui vont être importants pour leur vie symbiotique : les gènes *nod* (gènes de nodulation) ainsi que les gènes *nif* et *fix* (gènes de fixation d'azote). Ces gènes peuvent être portés sur le chromosome bactérien au niveau d'une région spécifique appelée îlot symbiotique ou sur un plasmide. Lorsque l'information est placée sur un plasmide (appelé alors plasmide symbiotique), elle peut facilement être transmise à

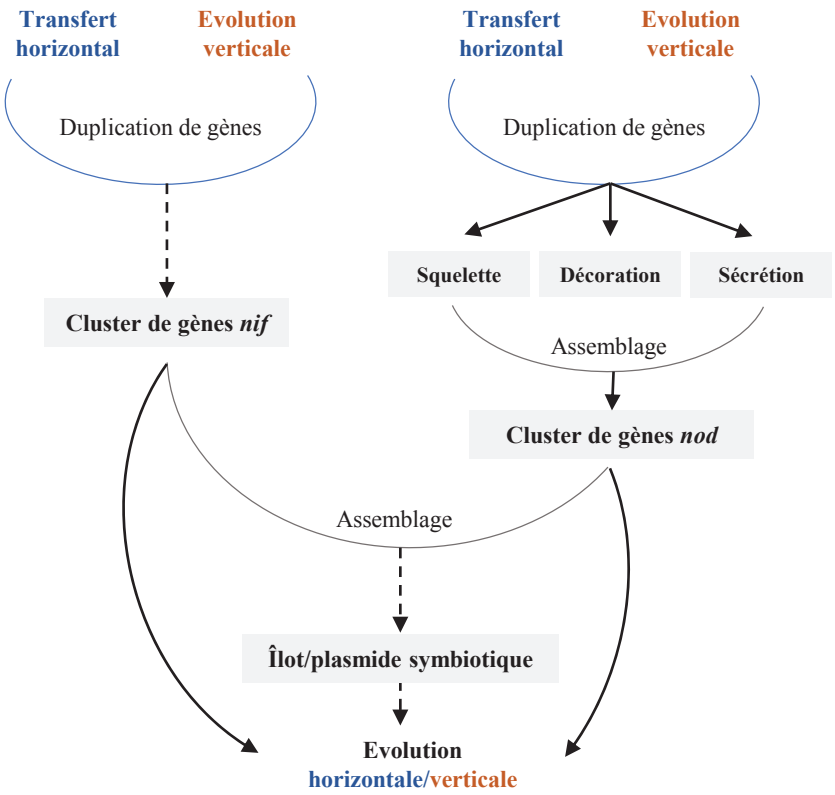


Figure 9. Modèle évolutif simplifié des gènes clés de la symbiose (adapté de Wang et al., 2019).
Les gènes impliqués dans la synthèse des facteurs Nod peuvent avoir des origines différentes et sont assemblés dans un groupe de gènes appelés *nod*. Les groupes de gènes *nod* et *nif* peuvent être assemblés ultérieurement dans un îlot symbiotique (dans le chromosome) ou dans un plasmide symbiotique. Les groupes de gènes *nod* et *nif* peuvent subir une évolution distincte ou similaire. L'ensemble de ces acquisition est modulé par le transfert horizontal ou vertical des gènes symbiotiques.

d'autres rhizobiums par transfert horizontal et/ou latéral, souvent contrôlé par le *quorum sensing* ou même par les plantes (Tun-Garrido et al., 2003; Ling et al., 2016).

En parallèle des gènes *nif* (cf. 1.2.1), les gènes *fix* sont des gènes additionnels néanmoins nécessaires à la fixation d'azote en participant entre autres aux transferts d'électrons lors du cycle catalytique de la nitrogénase mais aussi à l'activation de certains gènes *nif* (David et al., 1988; Lee et al., 2000). Enfin, les gènes *nod* souvent retrouvés à proximité des gènes *nif*, codent pour des protéines impliquées dans la synthèse et la sécrétion de molécules signales appelées facteurs Nod (FNs). Ceux-ci ont longtemps été considérés comme indispensables à l'interaction symbiotique car leur mutation empêche le développement des nodules chez la majorité des légumineuses (Long, 1989). Cependant en 2007, il a pu être montré que des souches de *Bradyrhizobium* photosynthétiques sont capable de noduler des plantes du genre *Aeschynomene* spp. alors qu'elles sont dépourvues de gènes *nod* (Giraud et al., 2007). Cette particularité symbiotique sera plus longuement développée dans la partie 5.3 de cette introduction. Cependant malgré cette exception, les gènes *nod* restent un marqueur moléculaire majeur pour identifier des nouvelles souches de rhizobiums en addition des gènes *nif*.

2.1.3. Un double mode de vie : libre ou symbiotique

Bien que les rhizobiums soient désignés ainsi pour leur capacité à s'associer symbiotiquement avec les légumineuses, ces bactéries peuvent aussi vivre en tant que saprophytes au sein de communautés microbiennes et en adoptant un mode de vie oligotrophe (Poole et al., 2018). Mais lorsque les bactéries se retrouvent dans le nodule d'une plante hôte, elles se différencient bien souvent en bactéroïdes endosymbiotiques et changent radicalement de mode de vie.

Ce double mode de vie, libre ou symbiotique, pourrait expliquer l'évolution de leurs génomes qui font partis des plus grands génomes bactériens (entre 5 à 10 Mb[6] en moyenne) (Poole et al., 2018). La plasticité de ces génomes qui codent pour de nombreux gènes nécessaires à la survie dans le sol mais aussi au sein d'une cellule végétale font des rhizobiums un groupe bactérien d'une grande variabilité génétique associée à un core-génome restreint (Amadou et al., 2008; Masson-Boivin et al., 2009). C'est grâce à l'acquisition des gènes symbiotiques que les rhizobiums ont pu, au cours de l'évolution, basculer d'un mode de vie saprophyte à celui de symbiotique (Sullivan & Ronson, 1998) (figure 9). Il a été ainsi montré que les symbiotes des légumineuses ont évolué par transfert horizontal de plasmides ou d'îlots génomiques importants pour l'interaction (Wang et al., 2019), ce processus étant favorisé par des flavonoïdes qui sont des molécules signales sécrétées par les végétaux. Ces données suggèrent donc que les plantes sont impliquées dans ce transfert de gènes et peuvent promouvoir les interactions symbiotiques entre légumineuses et rhizobiums (Ling et al., 2016). Cepen-

[6] Mégabase ; une mégabase représente une longueur d'un million de paires de bases.

Légumineuses à graines	**Production (Mt)**	***Bradyrhizobium* symbiote**
Soja	371,7	Oui
Arachide	53,9	Oui
Haricots	27,7	Non
Pois	32,9	Non
Pois chiche	15,8	Non
Lentilles	5,6	Non
Niébé	8,9	Oui
Pois d'Angole	5,5	Oui
Fève	5,9	Non
Haricot vert	1,3	Non
Lupin	1,4	Oui
Pois bambara	0,2	Oui
Autres	1	Non

Tableau 2. Rendement des légumineuses à graines d'intérêt agronomique en 2021 (fao.org/faostat).
En 2021, le rendement des légumineuses cultivées s'est élevé à 511,3 Mt dont 441,6 Mt ont pour origine un symbiote du genre *Bradyrhizobium*.

dant, ce passage à la vie symbiotique n'est sûrement pas l'objectif évolutif finale des rhizobiums au regard de plusieurs récentes analyses génomiques. Par exemple, lorsqu'on regarde la distribution des gènes symbiotiques *nif* et *nod* ou encore ceux codant pour le système de sécrétion de type 3 (T3SS[7]) chez les souches appartenant au genre *Bradyrhizobium*, il apparaît qu'il y a eu divers évènements d'acquisition et de perte de ces gènes au cours de l'évolution. Cela suggère que certaines souches ont abandonné leur vie symbiotique et que cette capacité à s'associer symbiotiquement avec les légumineuses n'est peut-être pas le style de vie dominant dans le genre *Bradyrhizobium* ou même chez les rhizobiums en général (Tao et al., 2021; Teulet et al., 2020).

2.1.4. Les *Bradyrhizobium* : le genre bactérien le plus abondant sur Terre

Le genre *Bradyrhizobium*[8] de la famille des Bradyrhizobiaceae, a été nommé en 1982 avec pour souche type *Bradyrhizobium japonicum* USDA6 anciennement nommé *Rhizobium japonicum* (Jordan, 1982). Il est l'un des vingt genres aujourd'hui identifiés comme contenant des rhizobiums. Une analyse récente des communautés bactériennes de sols provenant de 237 sites répartis sur les six continents a montré que les bactéries du genre *Bradyrhizobium* sont les plus abondantes et ubiquistes sur Terre (Delgado-Baquerizo et al., 2018). Le nombre d'espèces de *Bradyrhizobium* nouvellement identifiées ne cesse d'augmenter au fil des ans (http://edzna.ccg.unam.mx/rhizobial-taxonomy/) permettant une meilleure compréhension de ce genre bactérien aux propriétés notables. En effet, en plus d'être le genre le plus représenté dans les sols, c'est aussi celui qui est le plus utilisé en agriculture en raison de la capacité des *Bradyrhizobium* à interagir avec une vaste diversité de légumineuses tropicales ou tempérées dont certaines sont d'intérêt agronomique (tableau 2 – fao.org/faostat). Ainsi, sur la quantité totale de légumineuses produite en 2021, 86% ont pour symbiote des bactéries du genre *Bradyrhizobium*. Par exemple, différentes espèces de *Bradyrhizobium* sont les partenaires symbiotiques du soja, de l'arachide, du niébé ou encore du lupin (Xu et al., 1995; Barrera et al., 1997; Steenkamp et al., 2008).

On compte aujourd'hui 85 espèces différentes de *Bradyrhizobium* dont 70 ont été validées par le comité internationale sur la taxonomie des rhizobiums (Mousavi & Young, 2022), et c'est plus de 800 génomes annotés qui sont disponibles publiquement. Les génomes des souches de *Bradyrhizobium* sont généralement de grande taille par rapport aux autres taxons bactériens, allant de 6,8 Mb à 11,7 Mb (Avontuur et al., 2019) et sont associés à un pourcentage en GC élevé (≈60-65%). Par ailleurs, peu de souches portent des plasmides. A l'inverse des autres rhizobiums, où les gènes symbiotiques se

[7] Un des rôles du T3SS est de réprimer les mécanismes de défense de la plante pour favoriser l'infection symbiotique. Cet aspect-là est plus largement discuté dans la revue qui fait l'objet de la partie 6 de cette introduction.

[8] Du grec βραδύς signifiant "lent" et du terme générique rhizobium désignant les bactéries capables de noduler les légumineuses ; bien que tous les *Bradyrhizobium* ne sont pas symbiotiques (VanInsberghe et al., 2015).

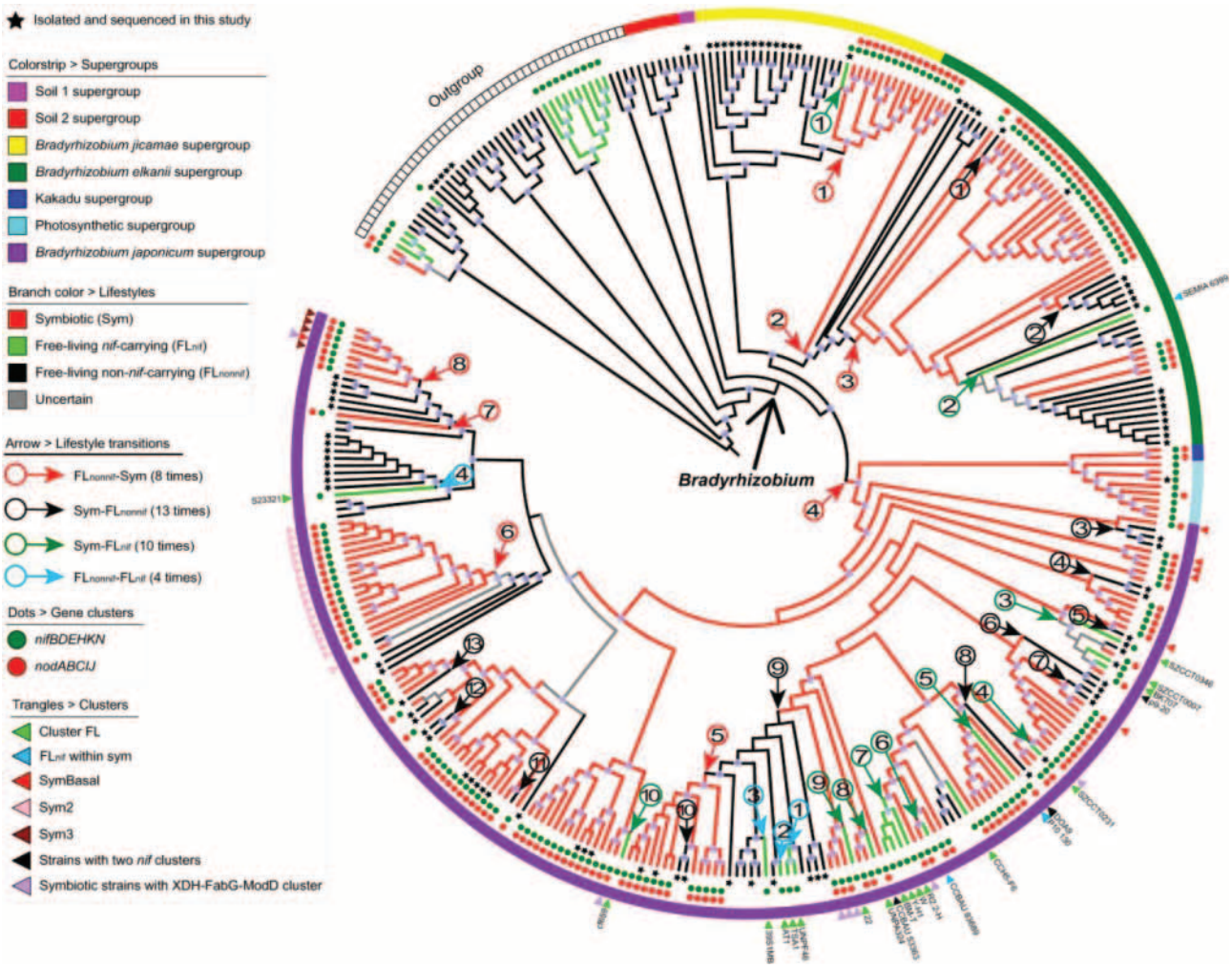


Figure 10. Arbre phylogénomique du genre *Bradyrhizobium* et histoire évolutive de son mode de vie (Tao et al., 2021).
Les différents supergroupes sont indiqués par les bandes colorées entourant l'arbre. Des souches de la famille Xanthobacteraceae constituent le groupe externe. Les cercles violet clairs indiquent les valeurs d'ultrafast bootstrap supérieures ou égales à 95% et calculées par IQ-Tree (Nguyen et al., 2015). Au cours de l'évolution, des *Bradyrhizobium* ont acquis la capacité à vivre à l'état symbiotique mais cette capacité a été aussi perdue par la suite chez certains.

situent souvent au niveau d'un plasmide symbiotique comme c'est le cas par exemple chez *Sinorhizobium meliloti* (Barnett et al., 2001), peu de *Bradyrhizobium* ont été rapporté comme porteur d'un plasmide symbiotique, comme chez la souche DOA9 (Okazaki et al., 2015). C'est donc sur le chromosome que les gènes *nod*, *nif* et *fix* sont dans la plupart des cas retrouvés[9], comme observé chez la souche *B. japonicum* USDA110 (Kaneko et al., 2002).

Phylogénie du genre *Bradyrhizobium*

De récentes analyses phylogénomique ont permis de mieux distinguer la diversité existante au sein du genre *Bradyrhizobium* (Avontuur et al., 2019; Garrido-Sanz et al., 2019; Teulet et al., 2020; Tao et al., 2021). Ces analyses montrent que ce genre est divisé en 7 clades plus ou moins larges, appelés aussi phylogroupes ou supergroupes (figure 10) :

- *Badyrhizobium japonicum* (189 souches).
- Photosynthétique (8 souches).
- Kakadu (2 souches).
- *Bradyrhizobium elkanii* (61 souches).
- *Bradyrhizobium jicamae* (32 souches).
- Soil 1 (2 souches).
- Soil 2 (7 souches).

Cependant, certains biais peuvent être admis concernant cette diversité observée puisqu'une majorité des souches utilisées pour ces analyses ont été isolées de nodules, souvent issus de légumineuses cultivées telle que le soja. L'isolation et le séquençage de nouvelles souches non issues de nodules devraient certainement changer l'architecture de cette phylogénie et élargir notre vision de ce genre bactérien

2.2. Le partenaire plante

2.2.1. Les légumineuses

La famille des Fabacées[10] est la troisième plus grande famille des angiospermes avec plus de 19 500 espèces après les Asteraceae et les Orchidaceae (Lewis et al., 2005; LPWG[11], 2013). Une récente classification phylogénique basée sur les séquences du gène *matK* de 91% des genres reconnus comme appartenant au groupe des légumineuses (698/765), a permis d'identifier six sous-familles au sein de la famille des Fabacées : les Caesalpinioideae, les Cercidoideae, les Detarioideae, les Dialioideae, les Duparquetioideae et les Papilionideae (LPWG, 2017; Zhang et al., 2020; Choi et al., 2022) (figure 11). A noter que la sous-famille traditionnellement reconnue sous le nom de Mimosoideae est aujourd'hui un clade distinct au sein des Caesalpinioideae et est désignée comme le clade mimosoidien en attendant une nouvelle classification des nouveaux Caesalpinioideae. Les études évolutives faites chez les légumineuses suggèrent qu'elles se sont diversifiées suite à l'extinction massive ayant eu lieu

[9] Sauf chez certains *Bradyrhizobium* photosynthétiques.
[10] Aussi appelée Fabaceae ou Leguminosae.
[11] The Legume Phylogeny Working Group.

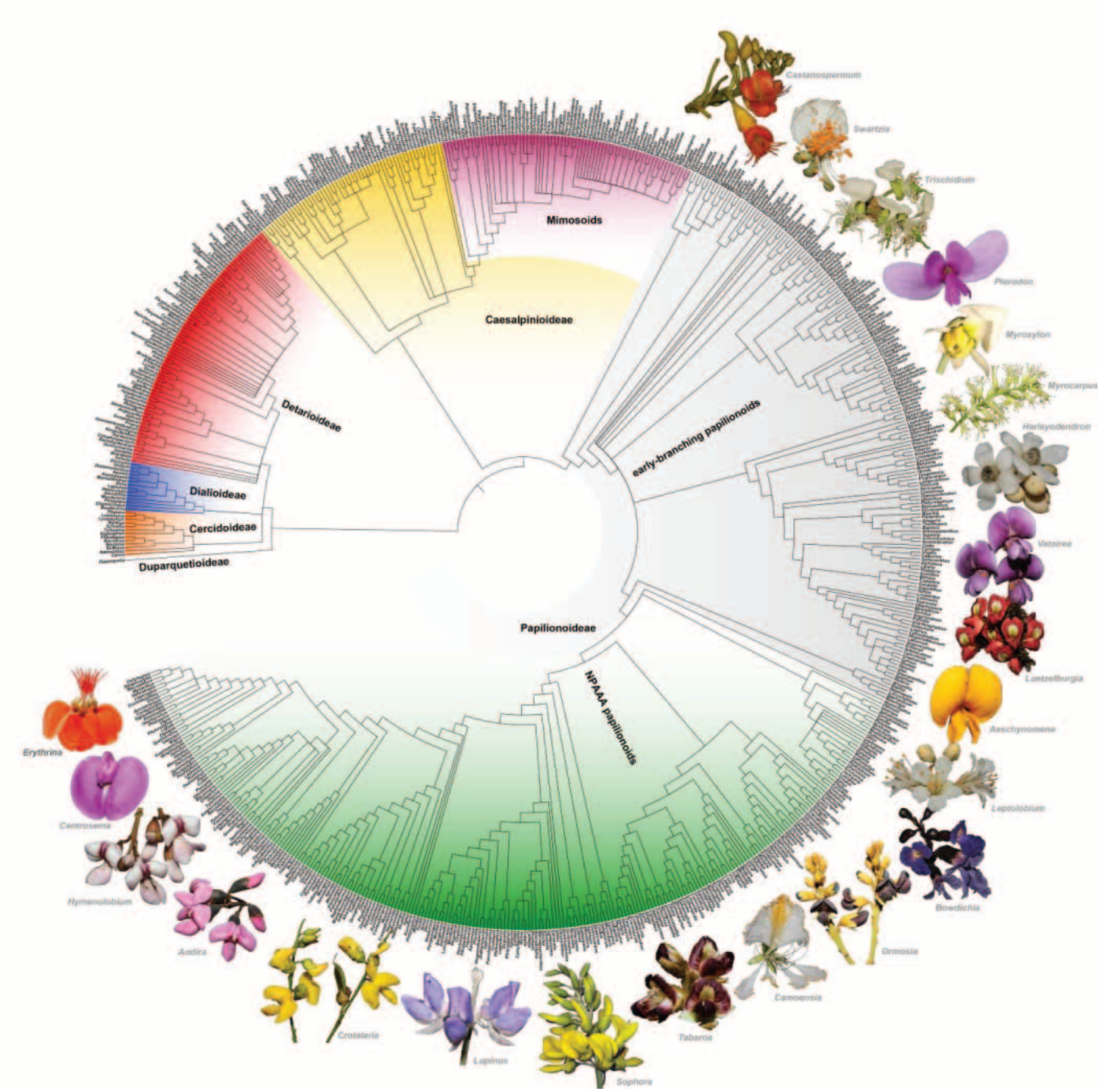


Figure 11. Phylogénie des légumineuses obtenue après assemblage du gène *matK* (Cardoso, 2023 – non publié).
Cet arbre consensus bayésien à règle majoritaire a été généré après avoir assemblé des données *matK* échantillonnées de manière exhaustive. Ces données proviennent principalement du LPWG (2017) avec des mises à jour supplémentaires incluant des séquences nouvellement publiées dans GenBank (Zhang et al., 2020 et Choi et al., 2022) ainsi que d'autres données non publiées. Pour le groupe des Papilionoideae, quelques fleurs des principales légumineuses de cette famille sont illustrées.

à la limite Crétacé-Paléogène et qui a vu l'extinction des dinosaures et la diversification des mammifères (≈ 66 millions d'années) (Koenen et al., 2021). Elles se sont ensuite diversifiées à travers des évènements de duplications entières de leur génome et à des phénomènes d'auto et d'allopolyploïdie (Centeno-González et al., 2021; Koenen et al., 2021). Enfin de récentes analyses transcriptomiques menées sur la légumineuse *Mimosa pudica* ont été obtenues en réponse à des souches de rhizobiums évoluées expérimentalement avec des compétences symbiotiques progressives. Cette étude a permis de démontrer que la réponse symbiotique était en place chez l'ancêtre commun plante de la symbiose rhizobium-légumineuse il y a plus de 90 millions d'années (Libourel et al., 2023).

2.2.2. Intérêts agronomiques, écologiques et économiques

Historiquement, la culture des légumineuses est un des meilleurs exemples de domestication des plantes par l'homme. De nombreuses espèces ont été adaptées pour l'agriculture et sont consommées dans le monde entier. Parmi elles, on peut par exemple citer le soja (*Glycine max*), les haricots (*Phaseolus* spp.), l'arachide (*Arachis hypogaea*), les lentilles (*Lens culinaris*), le pois chiche (*Cicer arietinum*) et les pois (*Pisum* spp.). Des espèces fourragères de climat tempéré ou tropical ont également été adaptées à la culture telles que la luzerne ou alfalfa (*Medicago sativa*), le trèfle (*Trifolium*) et la leucaena ou faux mimosa (*Leucaena leucocephala*) (LPWG, 2013).

Agronomiquement, les légumineuses sont d'un intérêt économique majeur juste après les Poacées, ou Poaceae/graminées, de par leur capacité inhérente à fixer l'azote atmosphérique améliorant la fertilité des sols (LPWG, 2017). Leur capacité à s'associer avec des rhizobiums pour la fixation d'azote symbiotique favorise donc une productivité agricole durable tout en enrichissant leur écosystème. Elles sont par ailleurs cultivées dans le monde entier et leur consommation participe à un régime alimentaire bénéfique pour la santé grâce notamment à la teneur en protéines élevée de leurs graines. Le manque d'intérêt porté sur l'utilisation des légumineuses dans l'agriculture par le passé, sûrement lié à des difficultés pour les cuisiner (temps de cuisson très importants pour certaines et la présence de substance anti-nutritionnelles pour d'autres), a conduit à une altération de la santé humaine, de la sécurité nutritionnelle mais aussi de la production alimentaire durable (Foyer et al., 2016). Il est donc primordial de porter un intérêt majeur sur la culture des légumineuses en raison des avantages multiples qu'elles peuvent apporter. Il existe ainsi un réel intérêt à remplacer dans l'alimentation humaine les protéines d'origine animale qui ont un coût écologique énorme par des protéines d'origines végétales issues des légumineuses. C'est entre autres pour cela que de nombreuses stratégies sont développées pour améliorer la culture des légumineuses et optimiser la sélection variétale, le rendement et la résistance aux stress biotiques et abiotiques (Foyer et al., 2016;

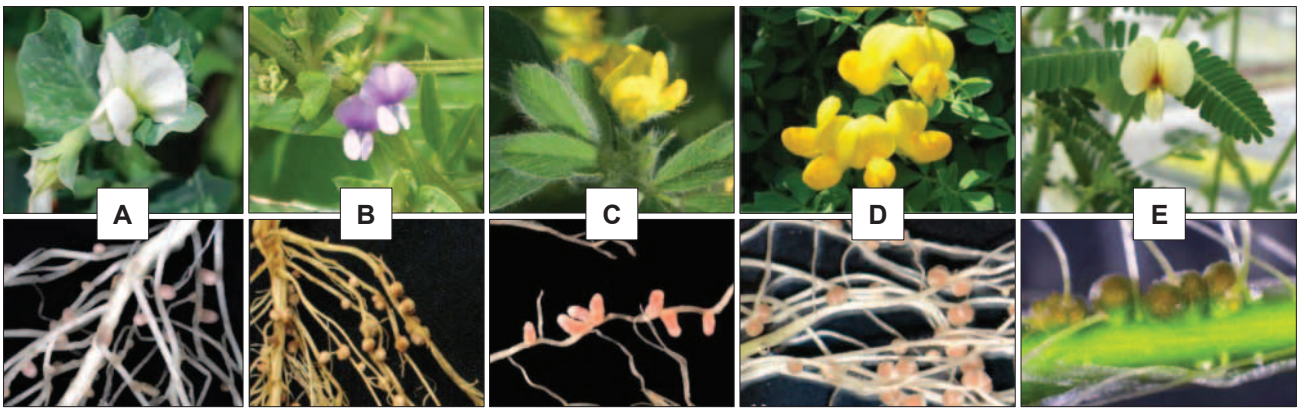


Figure 12. Fleurs et nodules racinaires des légumineuses modèles.
(**A**) Fleurs (Wikipedia) et nodules de *Pisum sativum* (Tsyganov & Tsyganova, 2020). (**B**) Fleurs (Wikipedia) et nodules de *Glycine max* (Bo Ren, Purdue University). (**C**) Fleurs (Eric Pensa) et nodules de *Medicago truncatula* (Wikipedia). (**D**) Fleurs (Humphreys, 2014) et nodules de *Lotus japonicus* (Suzaki et al., 2019). (**E**) Fleur (IRD – LSTM) et nodules d'*Aeschynomene evenia* (https://aeschynomenebase.fr). La racine apparaît verte car elle a été exposée à la lumière lors de sa croissance.

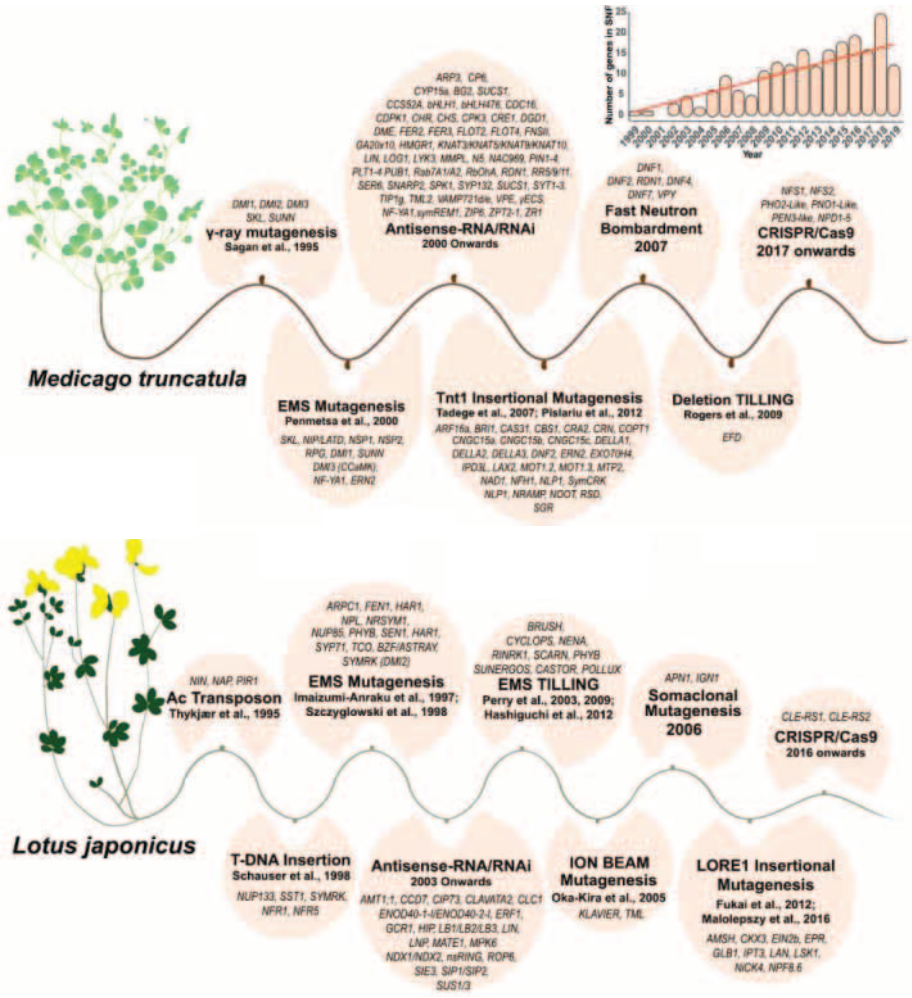


Figure 13. Chronologie de la découverte des gènes symbiotiques chez *M. truncatula* et *L. japonicus* (Roy et al., 2020).
L'obtention de nombreuses banques de mutants à partir d'une variété d'agents mutagènes chez les deux légumineuses modèles à permis depuis 1995 de découvrir de nombreux acteurs de la symbiose rhizobium/légumineuse. Le graphique en haut à droite montre le nombre de gènes découverts chaque année jusqu'en 2019. Certains de ces gènes seront détaillés dans les parties qui suivront cette introduction.

Raina et al., 2023). Par ailleurs certains arbres appartenant aux genres *Calliandra* (ex : pompon de marin rose), *Inga* (ex : pois doux), *Gliricidia* (ex : lilas étranger) ou encore les acacias (ex : *Acacia albida*[12]), peuvent s'associer avec des rhizobiums et représentent un intérêt majeur pour l'agroforesterie tropicale, la restauration des forêts et l'amélioration des sols (LPWG, 2013).

2.2.3. Les Papilionoideae : la plus large sous-famille des Fabacées

Avec environ 13800 espèces, les Papilionoideae[13] est la sous-famille la plus largement distribuée des Fabacées (Lewis et al., 2005). Leurs fleurs dont la forme évoque le plus souvent celle d'un papillon sont caractéristiques de cette sous-famille (figure 11 et 12). Celle-ci se divise en quatre clades : i) les Genistoïdes, ii) les Dalbergioïdes dont fait partie le genre *Aeschynomene* (cf. 5.2) mais aussi l'arachide *Arachis hypogaea*, iii) les Hologalegina où se placent le pois ou encore *Sesbania rostrata* capable de nodulation de tige et iv) les Phaseoloïdes/Milletioïdes dont font partie les deux grandes espèces modèles *Lotus japonicus* et *Medicago truncatula* (Doyle & Luckow, 2003; LPWG, 2013). A noter que c'est chez les Papilionoideae et les Mimosoideae que la nodulation par les rhizobiums est la plus courante avec plus de 97 % d'espèces nodulées (Sprent, 1995). Malgré cette large diversité au sein des légumineuses, seules quelques espèces d'intérêt agronomique (tels que le pois *Pisum sativum* (Tsyganov & Tsyganova, 2020) et le soja *Glycine max*) et les légumineuses modèles *M. truncatula* et *L. japonicus* et plus récemment *Aeschynomene evenia*, ont été étudiées (figure 12 et 13). Ainsi les légumineuses modèles ne couvrent qu'une petite fraction de la diversité existante des légumineuses et sont pour la plupart relativement proches phylogénétiquement, ce qui limite finalement nos connaissances globales sur cette famille (Doyle & Luckow, 2003).

[12] Synonyme : *Faidherbia albida*.

[13] Aussi appelés Faboideae, Papilionacea ou IRLC clade pour "inverted repeat-lacking clade" car cette famille est caractérisée par la perte d'une des deux répétitions inversées de 25 kb dans le génome plastidique que l'on retrouve dans la plupart des plantes terrestres.

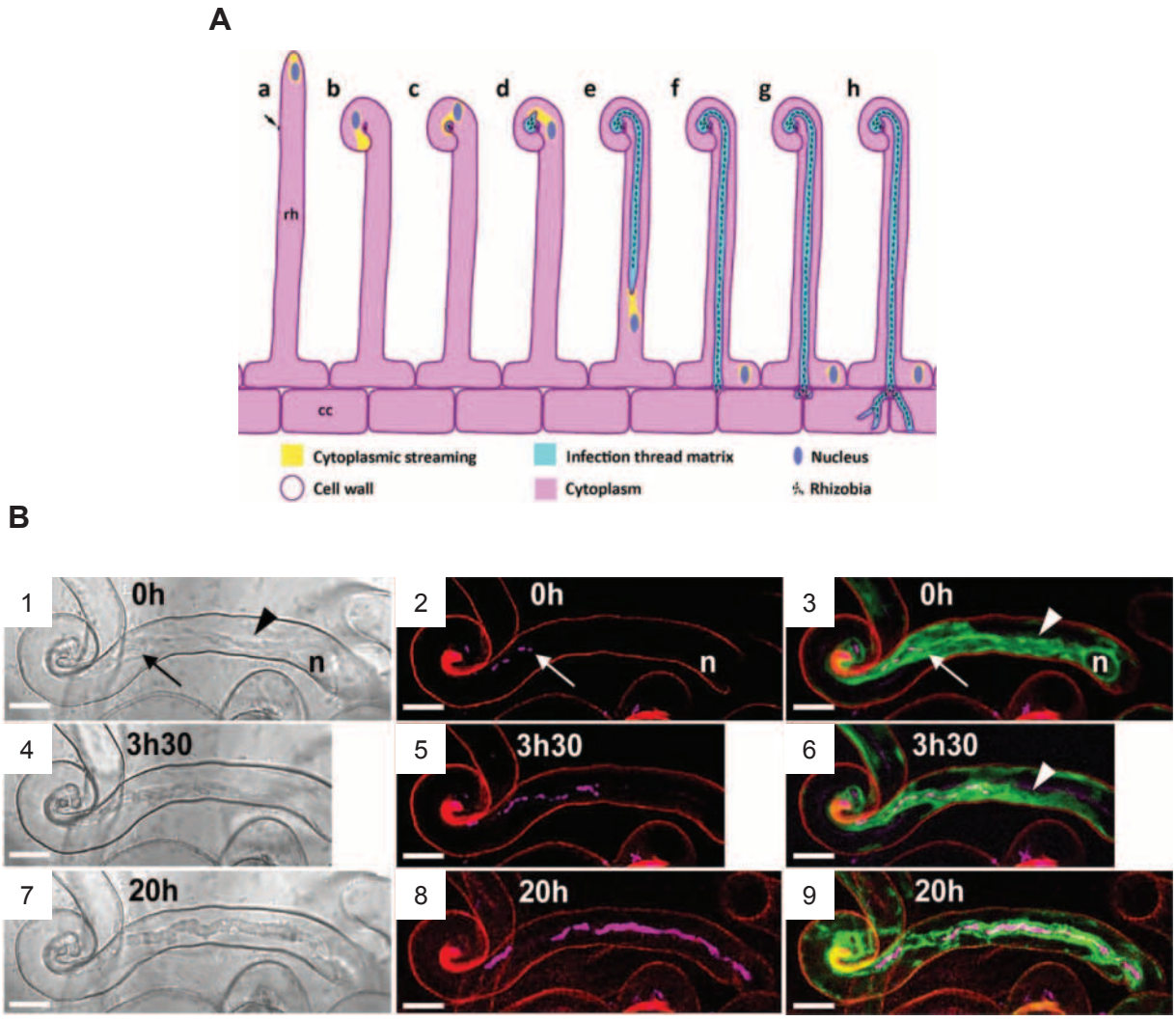


Figure 14. Infection intracellulaire par cordon d'infection.
(**A**) Représentation schématique de la formation d'un cordon d'infection, depuis l'adhésion du rhizobium au poil racinaire (rh) jusqu'à son arrivée dans une cellule corticale (cc). Contact entre un rhizobium et l'extrémité d'un poil racinaire (a) ; formation de la structure en crosse de berger créant une poche aussi appelée chambre d'infection (b) ; les bactéries commencent à se diviser ce qui provoque une courbure du poil absorbant qui emprisonne la bactérie dans une poche d'infection (c) ; la poche d'infection s'élargit à mesure que les bactéries se divisent et que le cordon d'infection se forme en face de la micro-colonie bactérienne (d) ; le cordon d'infection s'allonge et est guidé par le noyau (e) jusqu'à atteindre la base de la cellule (f) où de nouveaux cordons sont formés par ramification (g) pour *in fine* atteindre les cellules du cortex où les rhizobiums seront libérés (h) (Rae et al., 2021). (**B**) Progression du cordon d'infection dans un poil racinaire de *M. truncatula* exprimant le transgène 35S-GFP-HDEL (protéine verte associée au réticulum endoplasmique-RE) 4 jours après inoculation avec *S. meliloti* (fluorescence cCFP magenta). Les images en lumière blanche sont présentées à gauche (1, 4 et 7). Les images confocales correspondantes montrant les rhizobiums sont présentées en 2, 5 et 8 et combinées avec la GFP ciblée sur le RE en 3, 6 et 9. L'autofluorescence de la paroi végétale est visible en rouge. Un pont cytoplasmique riche en RE est indiqué avec une pointe de flèche (1 et 3) qui relie le noyau (n) au cordon d'infection en cours de développement (flèche). Durant la progression du cordon d'infection, les bactéries se multiplient à l'intérieur de ce nouveau compartiment (5 et 7). Les bactéries sont entourées de cytoplasme et le noyau du poil racinaire s'est déplacé le long du poil racinaire et n'est plus visible (6). Plus tardivement, le cordon d'infection s'est développé encore plus loin jusqu'à atteindre la base de la cellule pour ensuite se développer dans les cellules corticales. Echelle = 10 µm (adaptée de Fournier et al., 2008).

3. La symbiose nodulaire : de l'infection de la plante par la bactérie à la fixation d'azote

La symbiose entre les rhizobiums et les légumineuses est un processus complexe qui fait intervenir de nombreux acteurs moléculaires. Pour mieux comprendre son fonctionnement et identifier ses acteurs, des approches de génétiques classiques et inverses ont permis de découvrir près de 200 gènes chez la plante requis pour la symbiose fixatrice d'azote (figure 13). Ces travaux sont menés chez les légumineuses modèles et notamment chez les couples *Lotus japonicus*/*Mesorhizobium loti* et *Medicago truncatula*/*Sinorhizobium meliloti* (Roy et al., 2020). Dans les parties qui vont suivre, nous allons décrire les différentes étapes développementales et moléculaires qui surviennent lors de la reconnaissance de la bactérie par la plante jusqu'à la fixation d'azote et qui ont été décrites jusqu'à présent.

3.1. Différents modes d'infection

3.1.1. Infection intracellulaire par cordon d'infection

Une des stratégies d'infection utilisée en réponse aux rhizobiums pour pénétrer les tissus de la plante est l'infection intracellulaire par cordon d'infection (Fournier et al., 2008) (figure 14). Ce mode d'infection est celui observé chez le plus grand nombre de légumineuses (≈ 75%), à la fois tempérées et (sub)tropicales et est considéré comme étant le processus le plus avancé d'un point de vue évolutif (Madsen et al., 2010). Le processus d'infection débute lorsque les rhizobiums adhèrent à la surface du poil absorbant qui va s'enrouler alors en forme de "crosse de berger" et créer ainsi une poche qui devient une sorte de "piège" au sein duquel un rhizobium va former une micro-colonie en cours de division. Cette étape provoque conjointement un choc oxydatif médié par des espèces réactives à l'oxygène[14] et une dégradation de la paroi végétale. S'ensuit alors un remodelage de la membrane plasmique et un réarrangement du cytosquelette pour former un cordon d'infection (assimilable à un tube) qui va guider les bactéries à travers les cellules depuis la zone épidermale jusqu'au cortex où elles seront libérées dans des vésicules d'endocytose (Oldroyd et al., 2011). A noter que la déformation du poil requiert les récepteurs de la signalisation FN et que les FNs sont aussi impliqués dans le réarrangement du cytosquelette pour la déformation du poil racinaire (Cárdenas et al., 1998; Amor et al., 2003; Esseling et al., 2003). La bonne élongation et polarité du cordon sont médiées par un apport constant de vésicules (issues du Golgi et du réticulum endoplasmique) et par le cytosquelette (microtubules et microfilaments d'actine). Les rhizobiums jouent également un rôle important dans la formation et l'élongation du cordon d'infection par l'apport de composés sécrétés tels que les

[14] Aussi appelées ROS pour Reactive Oxygen Species.

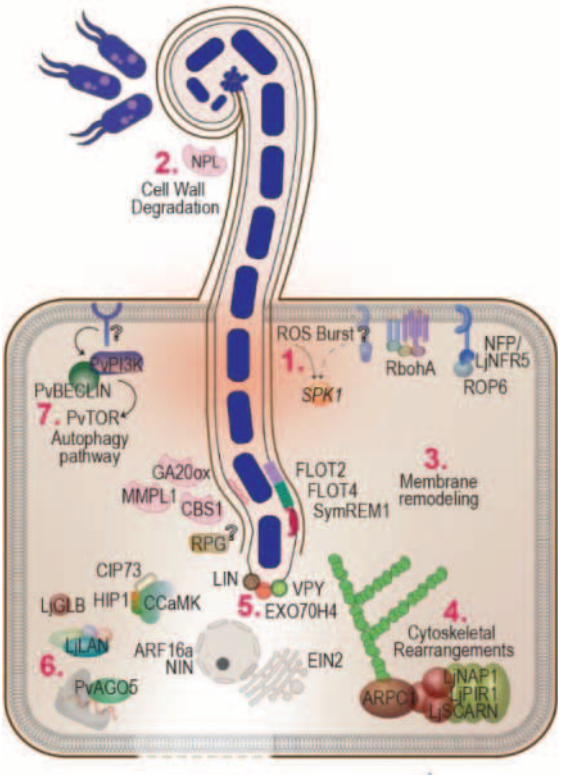


Figure 15. Représentation schématique des acteurs impliqués lors de l'infection par cordon d'infection (Roy et al., 2020). L'entrée des rhizobiums dans la cellule végétale à travers un cordon d'infection déclenche un choc oxydatif de ROS (1). Ce phénomène s'accompagne d'une dégradation des parois cellulaires (2), d'un remodelage membranaire (3) et d'un réarrangement du cytosquelette (4). Le cordon d'infection se dirige au sein de la cellule à l'aide d'un complexe situé à son extrémité : l'infectosome (5). En parallèle, des régulations transcriptionnelles et post-traductionnelles (6) ainsi qu'une activation de la voie de l'autophagie (7) assurent l'hébergement du symbiote dans la cellule végétale. L'ensemble des gènes présentés ici ont été identifiés chez les légumineuses modèles. Pour obtenir la majorité des références bibliographiques de chacun, se référer à la revue de Oldroyd et ses collaborateurs (2011). Cette liste n'est pas exhaustive puisque d'autres gènes ont été découverts depuis.

EPS[15] (Jones et al., 2007), créant une matrice fluide dans laquelle les bactéries peuvent se propager (Brewin, 2004) et jouant aussi un rôle important dans l'immunité (cf. 4.2). Les divisions simultanées des cellules du cortex racinaire et du péricycle déclenchent ensuite l'organogenèse du nodule lorsque l'infection est réussie.

Ce processus d'infection par cordon a pu être finement décrypté suite à l'étude de nombreux mutants déficients dans l'infection et dont les principaux gènes découverts sont retrouvés dans la figure 15 (Roy et al., 2020). Brièvement, un des premiers mutants découvert comme déficient pour l'infection par les rhizobiums code pour la protéine EIN2[16] cruciale pour la signalisation médiée par l'éthylène (Penmetsa & Cook, 1997). L'auxine a également été montrée comme jouant un rôle positif sur le nombre de cordons d'infection mis en place *via* le mutant ARF16[17] alors qu'à l'inverse les gibbérellines, à travers l'action des protéines DELLA[18], régulent négativement ce processus (Breakspear et al., 2014; McAdam et al., 2018). Ces observations soulignent l'importance du contrôle hormonal dans le développement et le nombre de cordons d'infection. L'importance de la réorganisation du cytosquelette dans le processus d'infection a pu être montré chez le lotier à travers par exemple le complexe SCAR/WAVE codant pour des protéines contrôlant l'assemblage du cytosquelette (Yokota et al., 2009). D'autres mutants affectés dans des pectinases ne sont plus capables de dégrader la paroi pectocellulosique de la plante ce qui stoppe l'entrée des rhizobiums dans la cellule (Xie et al., 2012). Les lipides, tels que les FLOTILLINs ont également un rôle important pour réguler la densité des récepteurs transmembranaires au niveau de la membrane plasmique permettant la bonne progression du cordon d'infection (Haney & Long, 2010). Plus récemment, un complexe exocyste (impliquant le trafic vésiculaire), appelé infectosome, a été montré requis pour assurer la bonne polarité du cordon d'infection et son maintien durant sa progression jusqu'au cortex (Lace et al., 2023).

3.1.2. Infection intercellulaire par crack-entry[19]

Ce second mode d'infection est moins répandu que le précédent et est retrouvé chez environ 25% des légumineuses (sub)tropicales (*Arachis* spp., *Sesbania* spp., *Stylosanthes* spp., *Neptunia* spp. et *Aeschynomene* spp.) (Boogerd & Van Rossum, 1997). Bien que ce type d'infection soit moins approfondi, des études faites chez *Sesbania rostrata*, *Arachis hypogaea* et *Aeschynomene* ont permis de mieux comprendre le fonctionnement de ce processus. L'infection intercellulaire nécessite une péné-

[15] Exo-polysaccharides de surface.
[16] Ethylene Insensitive 2.
[17] Auxin response factor 16.
[18] Les protéines DELLAs sont caractérisées par la présence d'un motif aspartate (D)-glutamate (E)-leucine (L)-leucine (L)-alanine (A).
[19] L'utilisation du terme crack entry provient de l'anglais "crack" qui signifie fissure et "entry" qui signifie entrée. C'est donc un mécanisme d'infection qui utilise des zones de fissures ou de brèche pour pénétrer à l'intérieur de la racine.

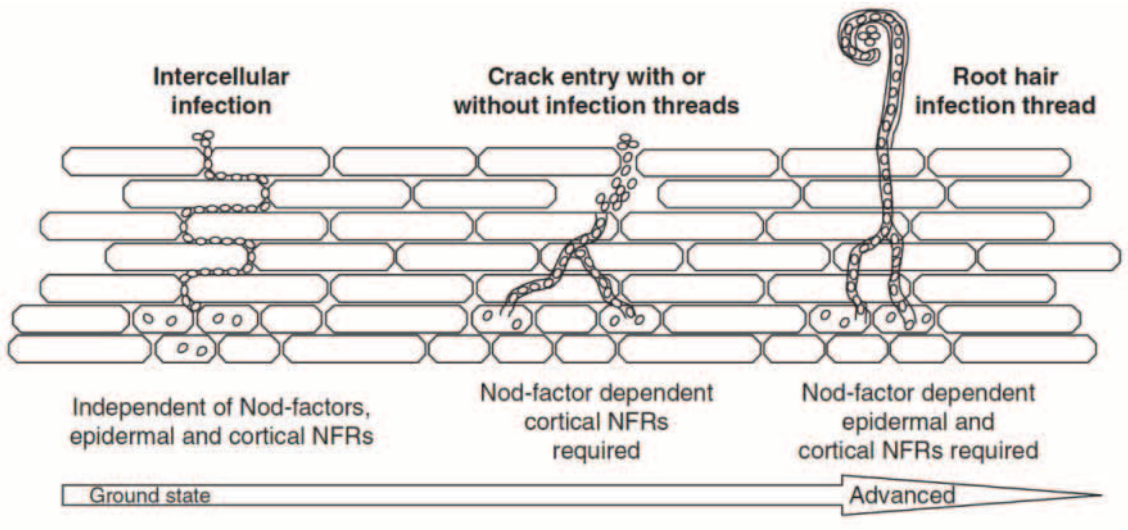


Figure 16. Différents modes d'infection : intercellulaire ou intracellulaire (*via* cordon d'infection).
Représentation schématique et modèle évolutif des différents modes d'infection intercellulaire ou intracellulaire lors de la symbiose rhizobium/légumineuse (Madsen et al., 2010). Un processus d'infection intercellulaire indépendant des facteurs Nod existe et est proposé comme ancestral. Ce mode d'infection sera discuté en partie 5.2. Le mode d'infection intercellulaire par crack-entry est retrouvé chez environ 25% des légumineuses et dépend des facteurs Nod uniquement pour la réponse corticale de la plante. Les bactéries pénètrent d'abord intercellulairement la racine au niveau de fissures avant de pénétrer dans le milieu intracellulaire en développant ou pas un cordon d'infection. Le mode d'infection intracellulaire est le plus répandu et est proposé comme le mécanisme d'infection le plus récent. Il est retrouvé chez les légumineuses modèles *L japonicus* et *M. truncatula* et dépend entièrement des facteurs Nod.

tration du rhizobium à travers l'épiderme. Pour cela la bactérie rentre initialement à travers des fissures provoquées par l'émergence des racines latérales (ou adventives si nodules de tige) (figure 16). Ce type de pénétration par crack-entry a aussi été observé chez la non-légumineuse *Parasponia* (van Velzen et al., 2018). Les rhizobiums se propagent ensuite dans les espaces intercellulaires de la racine. La bactérie va induire la mort de quelques cellules corticales pour former une poche d'infection intercellulaire et développer parfois un cordon d'infection (inter ou intracellulaire) (Madsen et al., 2010). Ce mode d'infection est retrouvé par exemple chez le couple *Sesbania/Azorhizobium* pour lequel il a été observé par ailleurs un mode d'infection intracellulaire par cordon d'infection selon les conditions de cultures (Goormachtig et al., 2004; Capoen et al., 2010). Comme pour l'infection intracellulaire, une dégradation des parois est observée ce qui favorise la propagation des rhizobiums jusqu'aux cellules corticales où a lieu l'internalisation des bactéries qui va conduire à la division des cellules du cortex pour former un primodium nodulaire (Sharma et al., 2020). Les nodules racinaires ainsi formés sont donc exclusivement retrouvés à la base d'une racine latérale où a eu lieu l'infection.

Contrairement au processus d'infection intracellulaire plus répandu, les poils racinaires des plantes infectées par crack-entry ne jouent pas de rôle précis dans l'infection. Par contre il est intéressant d'observer d'autres structures, qualifiées de poils axillaires, à l'émergence des racines latérales responsables parfois d'un gonflement épidermique pouvant être à l'origine de ces fissures et aidant à la colonisation bactérienne (Sharma et al., 2020; Quilbé, Montiel, et al., 2022). A l'inverse de l'infection intracellulaire, la stratégie crack-entry est moins ou pas dépendante de la synthèse des FNs pour les premières étapes de la colonisation (Herder et al., 2007). Chez les mutants dans les récepteurs des FNs, ce processus intercellulaires est environ 20 à 100 fois moins efficace que l'invasion normale par la courbure d'un poil racinaire chez *Lotus*, mais est néanmoins observé (Madsen et al., 2010).

L'infection intercellulaire est supposée ancestrale au regard des mécanismes sous-jacents impliqués dans la colonisation de la racine (Madsen et al., 2010). Cependant, étant donné que l'infection intracellulaire et intercellulaire se produisent chez des légumineuses apparues presque simultanément au cours de l'évolution, il est difficile de déterminer le mécanisme de colonisation ancestral (Koenen et al., 2020).

3.2. Mise en place de l'organe symbiotique : le nodule

3.2.1. Division cellulaire et organogenèse nodulaire

De façon concomitante à l'infection, la perception des FNs entraîne également une réactivation mitotique de certaines cellules corticales aboutissant à la formation d'un primordium nodulaire. Les divisions cellulaires ont lieu dans les cellules du cortex situées à la base de la zone d'infection et une

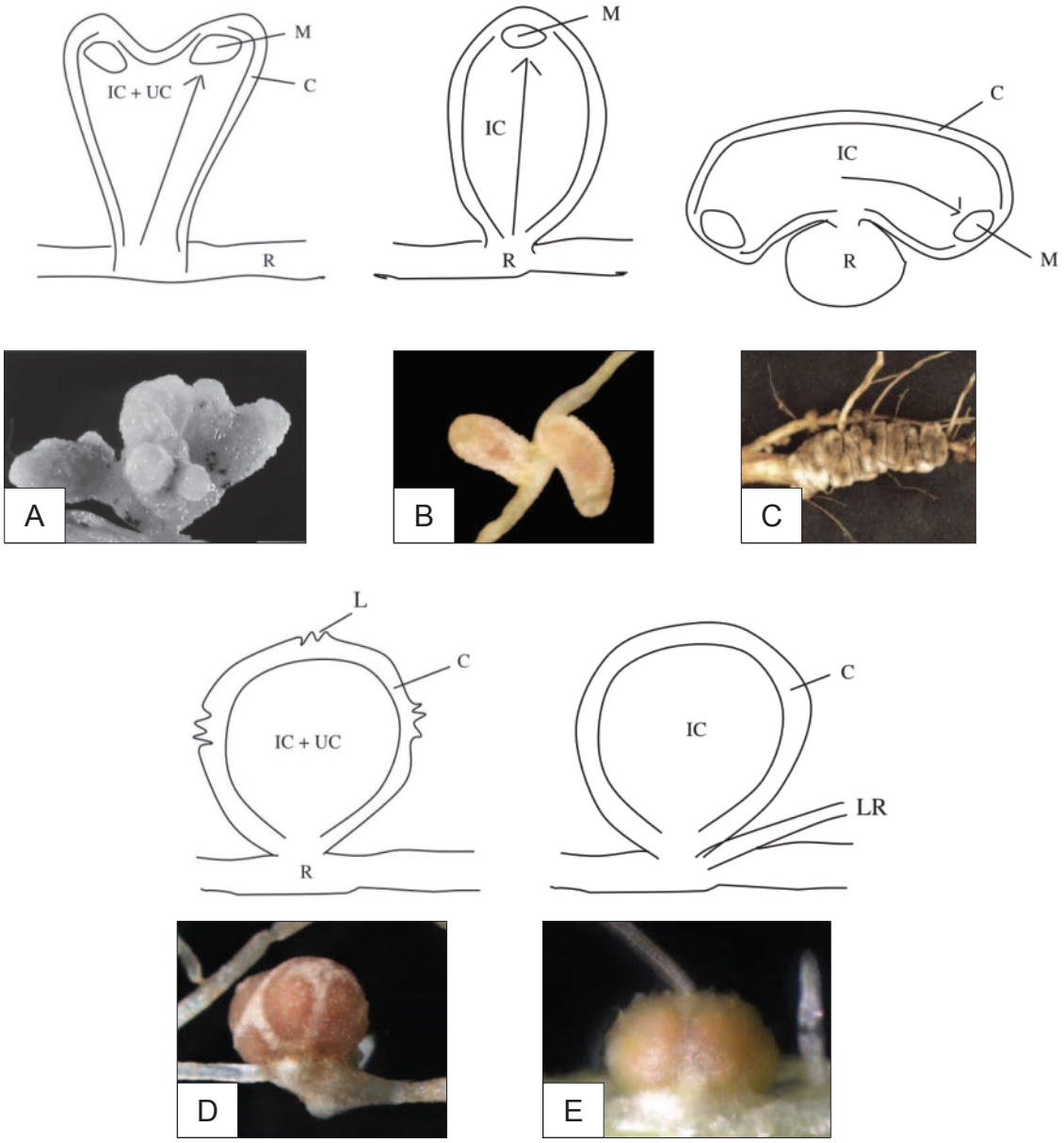


Figure 17. Morphologies de différents nodules rhizobiens (adapté de Sprent, 2007).
Principales caractéristiques des nodules de légumineuses observées en (**A**) *Crotalaria perrottetii* (Sy et al., 2001) ; (**B**) *M. truncatula* (Domonkos et al., 2013) ; (**C**) *Lupinus luteus* (Strozycki et al., 2007) ; (**D**) *Vigna mungo* (donnée personnelle) et (**E**) *A. indica* (donnée personnelle). IC, cellules infectées ; UC, cellules non infectées ; M, méristème ; C, cortex ; L, lenticelle ; R, racine secondaire ; LR, racine latérale. La flèche sur les schémas des nodules de type A, B et C (indéterminés) indique une gradation dans l'âge des cellules infectées, les plus jeunes étant vers la pointe de la flèche. La vascularisation du nodule est indiquée par un trait plein en périphérie de la zone infectée. A, B et C sont des nodules de type indéterminés, D un nodule de type déterminé et E un nodule de type aeschynoménoïde.

nouvelle vascularisation se met en place pour le nouvel organe en cours de formation : le nodule. Ces caractéristiques développementales sont communes à tous les nodules rhizobiens quel que soit le mode d'infection utilisé (Sy et al., 2001; Strozycki et al., 2007; Domonkos et al., 2013) (figure 17). Le nodule ainsi formé est constitué d'une zone centrale et d'une zone périphérique. Dans la première se trouve la zone infectieuse contenant les bactéries désormais capables de fixer biologiquement l'azote atmosphérique. Dans la seconde, trois tissus se distinguent avec chacun un rôle bien défini (Guinel, 2009). Dans la zone périphérique se trouve d'abord le tissu épidermique qui structure le nodule et assure son intégrité. Le tissu endodermique protège la zone centrale en formant une barrière de diffusion de l'oxygène nécessaire au bon fonctionnement de la nitrogénase. Enfin, le tissu vasculaire est le dernier tissu composant la zone périphérique et est au centre des échanges nutritionnels entre la plante et la bactérie. Selon le site d'initiation de ces divisions cellulaires et l'activité du méristème mis en place, différentes morphologies nodulaires peuvent être observées (Ferguson et al., 2010). Si les divisions sont initiées dans le cortex interne et le péricycle alors le nodule ainsi formé est de type indéterminé (Timmers et al., 1999). A l'inverse, un nodule est qualifié de déterminé s'il a pour origine une cellule du cortex moyen et externe (Hayashi et al., 2000). Une dernière morphologie de nodule déterminée dite de type aeschynoménoïde existe et sera discutée ultérieurement.

A l'œil nu, les nodules déterminés ont une forme plus ou moins sphérique et présentent souvent des lenticelles sous forme de bandes ou d'étoiles (figure 17D). Les nodules indéterminés sont quant à eux allongés et composés d'une ou plusieurs branches (figure 17A, B et C). Lorsqu'on observe plus attentivement l'intérieur des nodules on se rend alors compte qu'anatomiquement d'autres différences existent. Les nodules indéterminés possèdent un méristème toujours actif (ou persistant), présentent des cordons d'infection larges et ont des cellules infectées fortement vacuolisées. Les bactéries nouvellement différenciées en bactéroïdes pour fixer l'azote ont une durée de vie courte. Cette morphologie de nodule est le plus souvent retrouvée chez des légumineuses de climat tempéré telles que *Medicago truncatula* et le pois. A l'inverse, les nodules déterminés n'ont pas de méristème persistant, les cordons d'infections sont étroits, les cellules infectées sont très peu vacuolisées et les bactéroïdes ont une durée de vie plus longue que ceux retrouvés dans les nodules indéterminés. Les légumineuses associées à ce type de morphologie proviennent généralement de régions (sub)tropicales telles que *Lotus japonicus* et le soja.

En raison de l'existence d'un méristème apical produisant continuellement des nouvelles cellules qui deviendront par la suite infectées, il peut être observé chez les nodules de type indéterminé un gradient de développement (figure 18). Le nodule en constante élongation se divise alors en plusieurs zones (Ferguson et al., 2010) : i) une zone méristématique située à l'apex du nodule dépourvue de bactéries assurant une croissance en longueur continue, ii) une zone d'infection où les cellules corti-

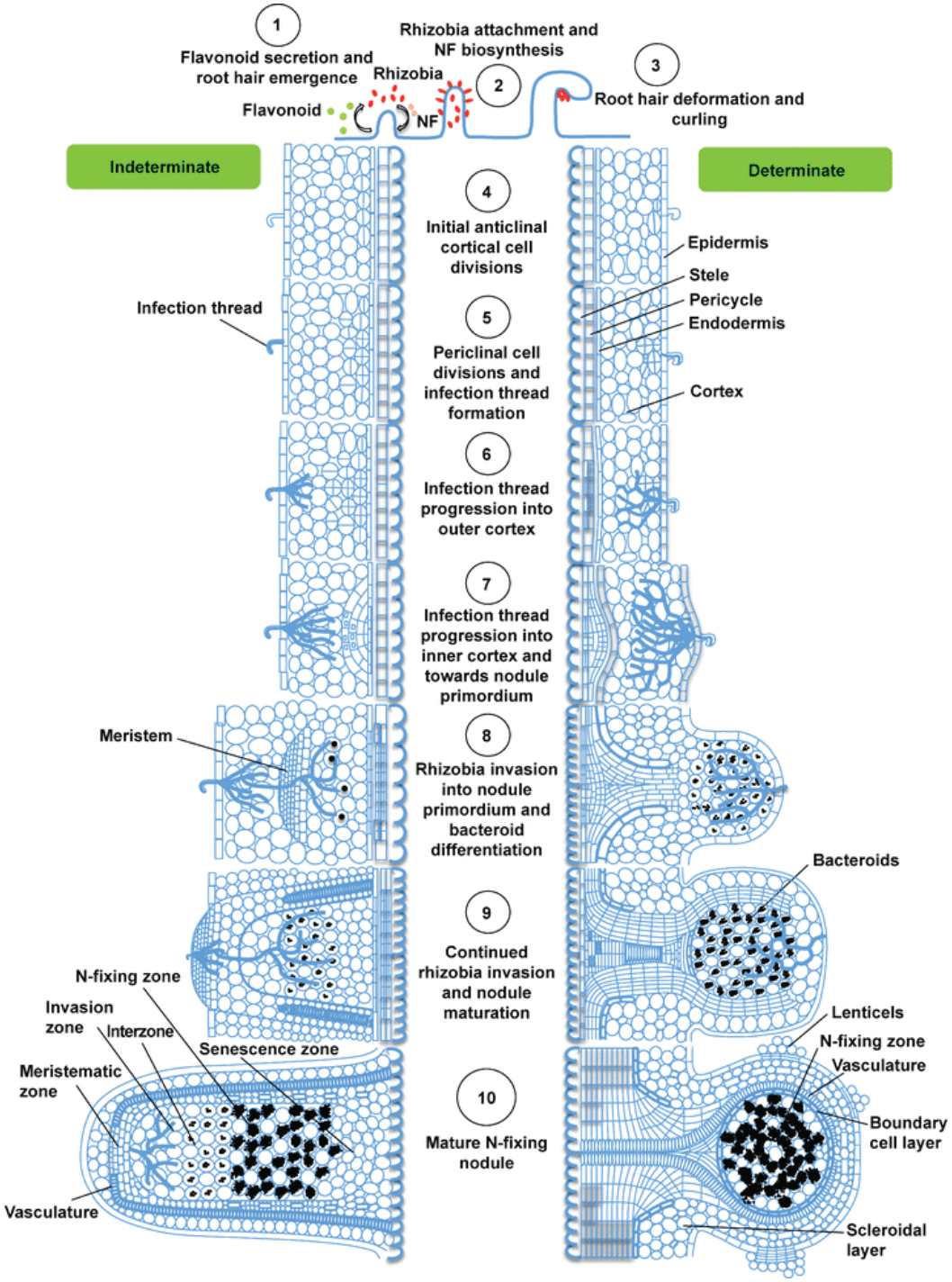


Figure 18. Comparaison schématique des étapes développementales entre un nodule de type indéterminé et un nodule de type déterminé (Ferguson et al., 2010).

Illustration des différentes étapes du développement d'un nodule de pois de type indéterminé (à gauche) et d'un nodule de soja de type déterminé (à droite). Ces étapes, numérotées de 1 à 10, s'initient au niveau des poils racinaires de la plante où les rhizobiums compatibles vont se retrouver piéger dans une poche d'infection formée par l'enroulement du poil en « crosse de berger » (1-3). Des cordons d'infections vont alors se développer et se propager jusqu'aux cellules du cortex où les divisions cellulaires vont initier le processus d'organogénèse nodulaire (4-7). Ces divisions cellulaires sont observées dans le cortex interne pour les nodules indéterminés et dans le cortex moyen ou externe pour les nodules déterminés. La libération des rhizobiums forme des symbiosomes dans lesquels ils se divisent en bactéroïdes fixateurs d'azote jusqu'à maturation du nodule (8-9). Chez les nodules de type indéterminé se développe un méristème apical persistant ce qui conduit à différentes zones successives d'invasion et de différenciation (10). A l'inverse, les nodules déterminés n'ont pas de méristème persistant et par conséquent toutes les cellules infectées ont un état de développement similaire (10).

cales sont envahies par les cordons d'infection du rhizobium qui sera par la suite internalisé au sein d'un compartiment nouvellement formé appelé symbiosome, iii) une interzone dans laquelle les bactéries se différencient en bactéroïdes pour débuter la fixation d'azote, iv) une zone de fixation où les bactéroïdes sont pleinement fonctionnels et v) une zone de sénescence lorsque les symbiosomes sont éliminés dans le cas de nodules âgés. A l'inverse, les nodules déterminés ont une croissance définie qui est stoppée lorsque le nodule est fonctionnel pour la fixation d'azote et ne possèdent donc pas de gradient de développement (figure 18). Cette activité du méristème explique la forme sphérique du nodule et non allongée comme chez les nodules de type indéterminé. La zone de fixation d'azote est la même que celle de l'infection qui correspond au tissu central du nodule. Lorsque les nodules deviennent sénescents, des nouveaux sont alors formés sur une portion plus jeune de la racine.

3.2.2. Contrôle moléculaire du développement nodulaire

Diverses hormones végétales coordonnent l'organogenèse nodulaire où les cytokinines et l'auxine jouent un rôle majeur (Ferguson & Mathesius, 2014; S. Roy et al., 2020). A travers les études faites sur le récepteur de cytokinines LHK1[20] chez *L. japonicus* (Tirichine et al., 2007) ou son orthologue CRE1[21] chez *M. truncatula* (Plet et al., 2011), il a pu être montré que cette hormone est nécessaire et suffisante pour induire la formation de nodules. Cette réponse aux cytokinines est restreinte au cortex et au péricycle (Plet et al., 2011) qui sont les assises cellulaires où se déroulent les divisions cellulaires durant le processus de nodulation. Plusieurs facteurs de transcriptions facilitent l'activation des cytokinines en réponse aux FNs pour initier l'organogenèse nodulaire. On peut entre-autre citer NIN[22] et NSP2[23] (Madsen et al., 2010) mais aussi EFD[24] (Vernié et al., 2008). Une conséquence directe de cette signalisation médiée par les cytokinines est une modulation de la réponse auxinique. Il a pu être montré que le transport polaire de l'auxine à travers les transporteurs PINs (transport d'efflux) est supprimé de manière locale au niveau de la zone infectée par les rhizobiums en faveur de la nodulation (Mathesius et al., 1998). Cette suppression passe par les cytokinines et conduit à une accumulation locale d'auxine permettant l'initiation des divisions cellulaires pour former ainsi un nouveau méristème nodulaire (Plet et al., 2011).

Le rôle d'une troisième hormone végétale dans l'organogenèse nodulaire a pu être identifiée à travers le mutant *sickle* insensible à l'éthylène chez *M. truncatula* (Prayitno et al., 2006). Chez ce mutant, un phénotype d'hypernodulation est observé qui est associé à une augmentation du transport

[20] Lotus Histidine Kinase 1.
[21] Cytokinin Response 1.
[22] Nodule Inception.
[23] Nodulation Signaling Pathway 2.
[24] Ethylene response factor required for nodule differentiation.

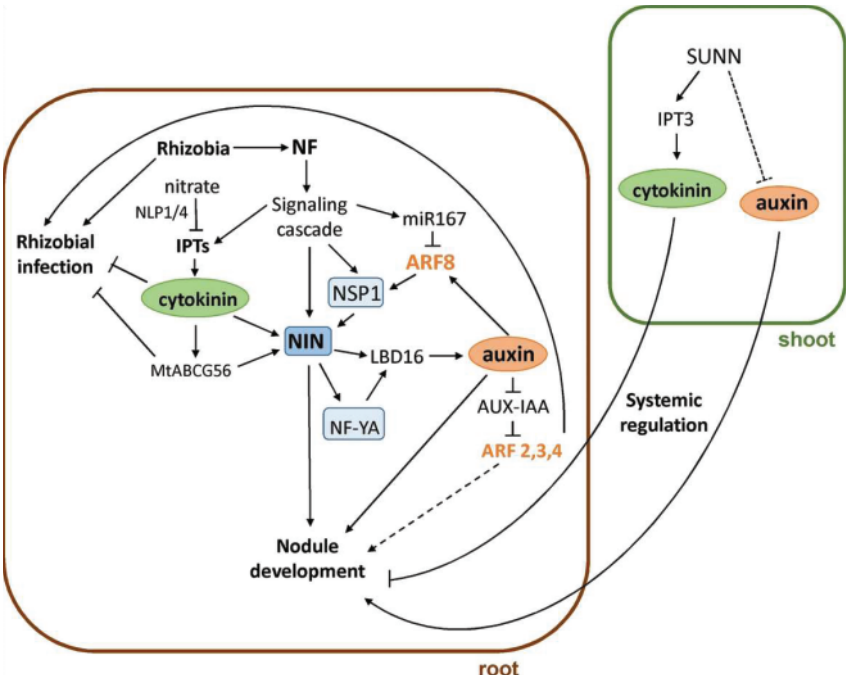


Figure 19. Régulation par l'auxine et les cytokinines durant l'organogénèse nodulaire (Azarakhsh et Lebedeva, 2023). Divers acteurs tels que des miRNA ou des facteurs de transcription régulent l'ensemble de cette signalisation hormonale et une communication entre les feuilles et les racines rentre également en jeu. Les traits en pointillés indiquent une action indirecte.

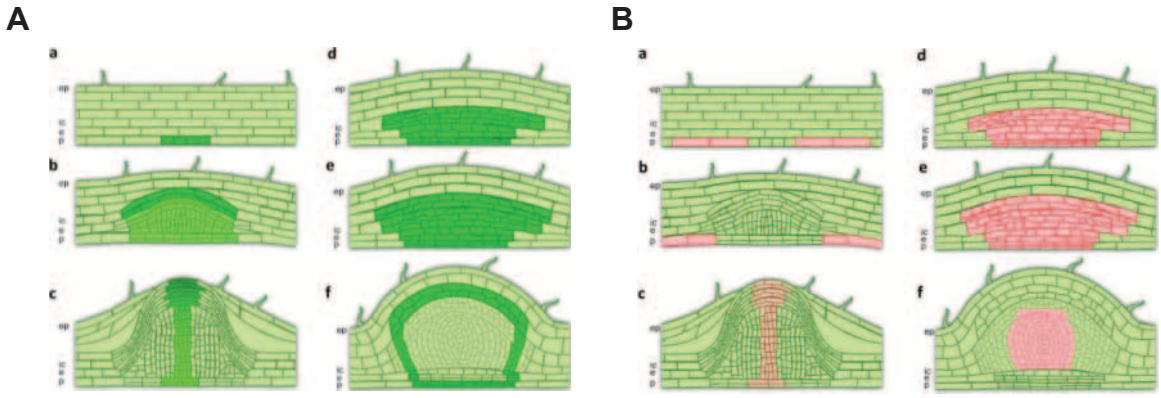


Figure 20. Réponse de l'auxine et des cytokinines durant le développement d'une racine latérale ou d'un nodule (Azarakhsh et Lebedeva, 2023).
(**A**) Réponse auxinique du tissu racinaire (vert foncé) pendant le développement d'une racine latérale (a-c) ou d'un nodule (d-f). Pour une racine latérale, ce sont les premières cellules du péricycle en division qui montrent une réponse à l'auxine (a). Plus tard, l'expression est observée dans le primordium racinaire en croissance (b) jusqu'à la formation d'un nouveau faisceau vasculaire central nouvellement formé (c). Pour un nodule, la réponse auxinique a lieu dans les cellules du péricycle mais aussi de l'endoderme et du cortex (d). Plus tard, cette réponse s'étend au cortex externe (e) puis une vascularisation périphérique se met en place (f). (**B**) Réponse des cytokinines du tissu racinaire (rose) pendant le développement d'une racine latérale (a-c) ou d'un nodule (d-f). Pour une racine latérale, la réponse aux cytokinines a lieu dans des cellules individuelles du pôle xylème du péricycle entre le jeune primordium en cours de développement (a). Puis la réponse aux cytokinines est réduite dans les cellules autour du primordium racinaire (b) pour ensuite avoir une réponse dans les cellules vasculaires (c). Lors du développement d'un nodule, le signal médié par les cytokinines est important au niveau du péricycle et des cellules corticales en cours de division (d et e). Lorsque le primordium nodulaire est formé, la réponse au cytokinine reste localisée au niveau des cellules centrales (f) (p péricycle; e endoderme; ic cortex interne; ep épiderme).

longue distance d'auxine depuis les feuilles jusqu'aux racines et à une augmentation du nombre de transporteurs PINs au niveau de la zone d'initiation du nodule (Prayitno et al., 2006). Il a aussi été montré que l'éthylène joue un rôle dans le positionnement des nodules en face des pôles du xylème par inhibition de la division des cellules corticales (Heidstra et al., 1997). Enfin, il a pu être montré que des cytokinines et des molécules liés aux cytokinines peuvent aussi être produites par la bactérie à travers la synthèse de purines (Okazaki et al., 2007). Le rôle de cette hormone a été étudié en profondeur chez le couple symbiotique *Aeschynomene*/*Bradyrhizobium* où elle contribue positivement à l'organogenèse du nodule mais n'est cependant pas essentielle à l'interaction (Podlešáková et al., 2013). D'autres acteurs ont depuis été identifiés à la fois dans le développement racinaire et le développement nodulaire et sont résumés dans la figure 19 (Azarakhsh & Lebedeva, 2023). L'ensemble de ces observations suggèrent que les plantes et les microorganismes ont co-évolués pour adapter leur développement à la symbiose (Desbrosses & Stougaard, 2011).

Un nodule est-il une racine modifiée ?

Des similarités entre le développement d'un nodule et celui d'une racine latérale peuvent être observées. En effet, les mêmes hormones sont impliquées à la fois dans le processus de nodulation et dans celui du développement racinaire (auxine, cytokinine et éthylène). Tout comme les racines latérales, les nodules sont des organes racinaires latéraux qui influencent l'architecture du système racinaire de la plante. Il y a une apparente coordination entre le développement racinaire et nodulaire, cependant l'origine cellulaire d'un nodule et d'une racine est différente (figure 20).

La mise en place d'un nodule nécessite le recrutement de facteurs contribuant au développement d'une racine latérale tels que les gènes PLETHORA[1] ou LBD16[2] (Soyano et al., 2021). Le positionnement du nodule ou de la racine latérale sur la racine principale est similaire, cependant le système vasculaire est différent : il est central pour une racine latérale et périphérique pour un nodule. De plus, c'est la bactérie qui contrôle le programme de développement du nodule et si elle meurt au cours du processus, alors la racine va reprendre son développement classique de racine latérale (Ferraioli et al., 2004). Les micro-organismes sont donc capables de modifier les voies de développement de l'hôte en faveur de la symbiose.

Les changements observés dans les niveaux d'auxine et de cytokinines pour la formation du méristème nodulaire peuvent aussi rappeler ceux observés pour la formation des méristèmes pour le développement des racines latérales. Cependant on relève un rôle inversé de ces hormones selon le développement mis en place : ce sont les cytokinines qui sont les premiers activateurs du méristème nodulaire alors que pour la formation d'une racine latérale, la signalisation associée aux cytokinines est éteinte pour la bonne formation du méristème racinaire. La suppression du signal médié par les cytokinines est donc un indicateur des premiers stades de la formation des racines latérales (Benková & Bielach, 2010). Ainsi, la balance hormonale dans une racine suite à la reconnaissance des rhizobiums est un équilibre qui peut favoriser l'initiation des nodules mais supprimer l'initiation des racines latérales. On peut donc dire qu'un nodule est une racine partiellement modifiée.

[1] Les gènes PLETHORA ont été initialement découverts chez *Arabidopsis thaliana* et sont transcrits en réponse à l'accumulation d'auxine. Ils médient la structuration de la niche des cellules souches de la racine d'*Arabidopsis*.

[2] Lateral Organ Boundaries-Domain protein 16, le domaine LOB étant un régulateur clé du développement chez les plantes.

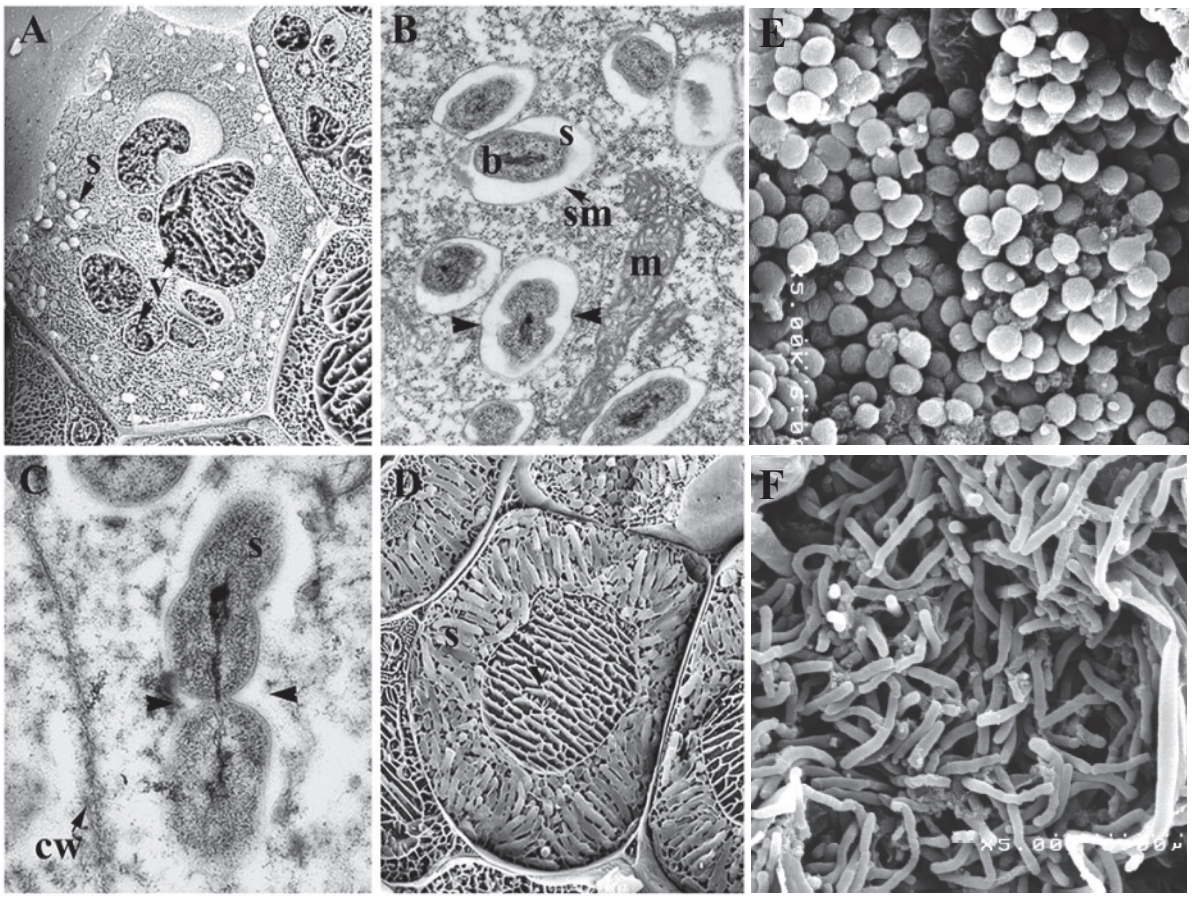


Figure 21. Cellules infectées de nodules observées chez diverses légumineuses (Coba de la Peña et al., 2018; Bonaldi et al., 2011).
(**A**) Quelques symbiosomes sont visibles au sein d'un jeune nodule de *Medicago sativa* où la vacuole se désintègre. (**B-C**) Symbiosomes en cours de division (indiqué par des pointes de flèche) chez des jeunes nodules de *M. sativa* (B) ou *Lupinus albus* (C). (**D**) Cellules infectées matures abritant des symbiosomes matures également (*M. sativa*). (**E**) Intérieur d'une cellule infectée contenant des bactéroïdes de forme sphérique chez *A. indica*. (**F**) Intérieur d'une cellule infectée contenant des bactéroïdes de forme allongée chez *A. afraspera*. Les images ont été prises au microscope électronique à balayage (A, D, E, F) ou à transmission (B, C). b, bactéroïde ; cw, paroi cellulaire ; m, mitochondries ; s, symbiosome ; sm, membrane du symbiosome ; v, vacuole.

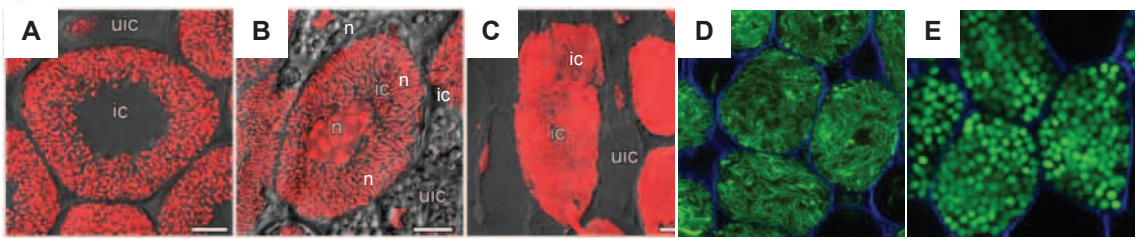


Figure 22. Différences morphologiques des bactéroïdes (adapté de Bonaldi et al., 2011 et Kitaeva et al., 2021).
Cellules infectées par des rhizobiums compatibles chez 5 espèces de légumineuses différentes. (**A-C**) Coloration de l'ADN au iodure de propidium qui colore en rouge les noyaux et les bactéries. (**D-E**) Souche ORS285 marquée à la GFP. On observe des morphologies variables des bactéroïdes : type allongé/branché (A), allongé (B et D), gonflé (C) ou sphérique (E). n, noyau ; ic, cellule infectée ; uic, cellule non infectée. Echelle : 10 µm.

3.2.3. Internalisation des bactéries au sein du symbiosome et différenciation en bactéroïdes

Lorsque l'infection par les rhizobiums est réussie, un contact direct entre la bactérie et la membrane végétale dénuée de paroi entraîne la libération des bactéries par endocytose. Initialement vivant à l'état libre, les rhizobiums en infectant la plante se retrouvent internalisés en micro-colonies entourées d'une membrane appelée péribactéroïdienne formant un nouveau compartiment cellulaire : le symbiosome (figure 21). S'ensuit le développement sphérique ou allongé du nodule selon la légumineuse infectée. Les bactéries vont alors devoir s'acclimater à leur nouvel environnement intracellulaire et mettre en place la fixation d'azote atmosphérique. Pour cela, elles vont se différencier en bactéroïde qui est la forme active capable de fixer le N_2.

Au sein des nodules de type déterminé comme ceux de *L. japonicus*, les symbiosomes contiennent généralement deux ou plusieurs bactéroïdes de taille similaire à celle des bactéries à l'état libre. A l'inverse, les symbiosomes des nodules indéterminés comme ceux de *M. truncatula*, contiennent un seul bactéroïde pléomorphe[25] (Oldroyd et al., 2011). En effet, les bactéroïdes provenant de ces nodules subissent une endoréplication[26] de leur génome pouvant aller jusqu'à 24 chromosomes contre 1 à 2 pour des bactéries libres (Mergaert et al., 2006). Ces mêmes bactéroïdes voient leur membrane devenir perméable à l'inverse de leur forme libre et leur différenciation est dite "terminale " car les bactéries perdent la capacité de se multiplier et de revenir à un état de vie libre initial. Ces changements génétiques résultent en des changements morphologiques des bactéroïdes qui peuvent être gonflés, allongés, sphériques et allongés ou enfin ramifiés (Montiel et al., 2017; Kitaeva et al., 2021) (figure 22).

L'ensemble de ces morphologies est hôte-dépendant et il a été montré que les légumineuses conduisant à une différenciation terminale appartiennent au clade des Papilionoideae des légumineuses.

Le fait qu'une même bactérie puisse avoir des morphologies différentes selon la légumineuse infectée renforce l'idée que le partenaire plante joue un rôle crucial dans cette différenciation. En effet, pour une même souche de *Bradyrhizobium*, diverses morphologies sont observées chez différentes espèces appartenant au genre *Aeschynomene* (Bonaldi et al., 2011; Lamouche et al., 2019). Chez les autres légumineuses, il a été observé que les bactéroïdes provenant des nodules tels que ceux de *L. japonicus* contiennent des bactéries où la différenciation n'est pas terminale (Müller et al., 2001). Leur membrane reste imperméable et il n'y a pas d'endoréplication observée (Mergaert et al., 2006). L'ensemble de ces données ont permis de conclure que la différenciation des bactéroïdes et

[25] Du grec ancien πλείων, *pleiôn* ("plusieurs") et μορφή, *morphê* ("forme"). Se dit d'un microorganisme qui possède la propriété de changer de forme, de fonction(s) biologique(s) et de mode de reproduction en réponse à diverses conditions environnementales.

[26] Continuité de cycles cellulaires qui ont lieu sans division cellulaire associée. Cela conduit à une multiplication du nombre de copies du génome au sein d'une même cellule.

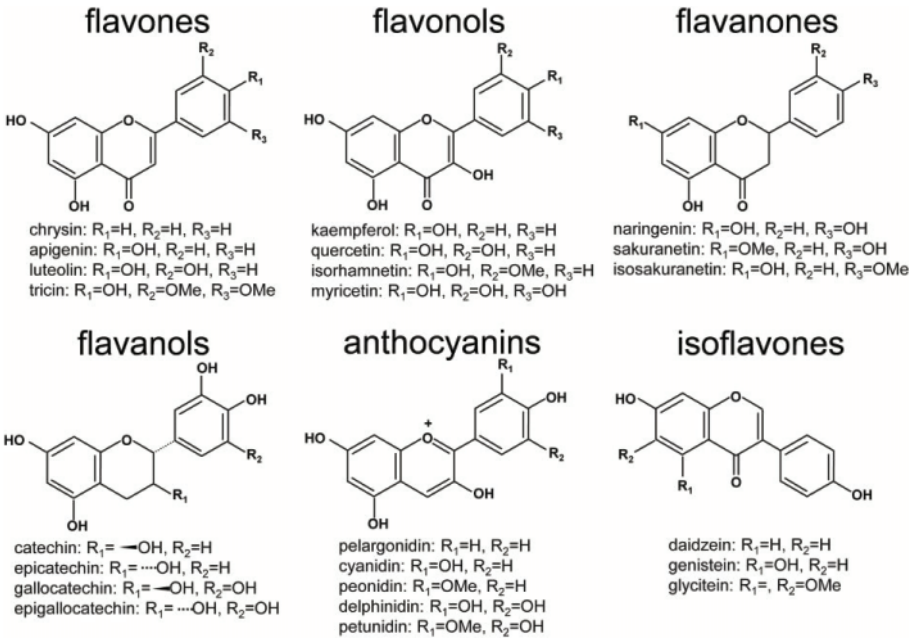


Figure 23. Structure des principaux composés organiques constituant les flavonoïdes (Tohge et al., 2017).
Les flavonoïdes peuvent être divisés en 6 sous-classes principales et quelques exemples sont donnés pour chacune des sous-classes existantes. Cette grande diversité des flavonoïdes produits par les plantes implique des rôles multiples au sein de la rhizosphère. Chez les légumineuses, la sécrétion de flavonoïdes active chez les rhizobiums l'expression des gènes *nod* à travers leur reconnaissance par la protéine NodD. La bactérie va par la suite initier les premières étapes du programme de nodulation.

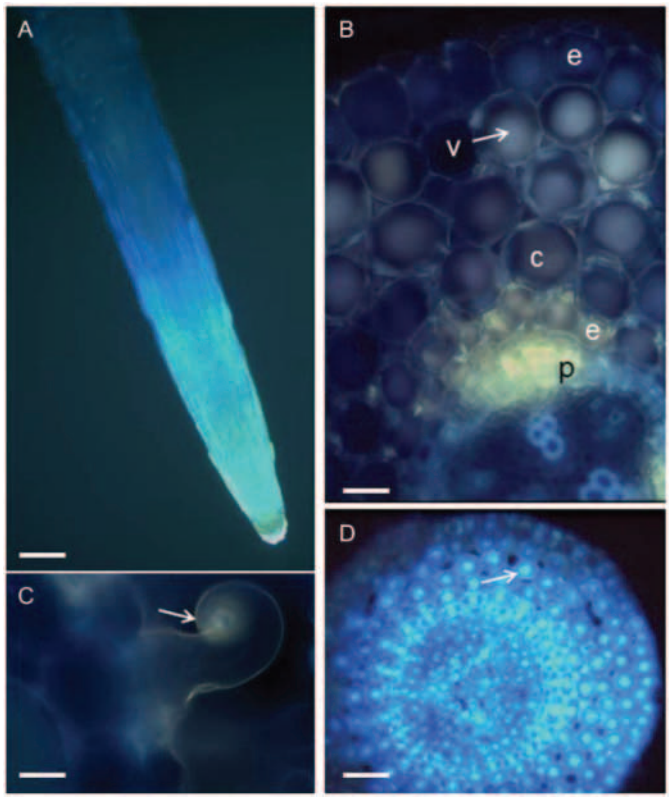


Figure 24. Visualisation de l'accumulation de flavonoïdes dans et entre les cellules racinaires (Hassan & Mathesius, 2012).
(**A**) Accumulation de flavonoïdes visibles par fluorescence à l'extrémité d'une racine de *M. truncatula*. (**B**) Des flavonoïdes spécifiques (fluorescence de différentes couleurs) s'accumulent dans différents types de cellules du trèfle blanc (*Trifolium repens*). (**C**) Flavonoïdes (fluorescence orangée) localisés dans la paroi cellulaire et/ou la membrane d'un poil racinaire en « crosse de berger » de *M. truncatula* suite à l'inoculation avec *S. meliloti*. (**D**) Localisation nucléaire des flavonoïdes dans une coupe transversale de l'extrémité d'une racine de trèfle blanc. Sous excitation UV (350 nm), il est possible de visualiser la fluorescence des flavonoïdes après coloration avec de l'acide diphénylborique-2-aminoéthyl ester ; p = péricycle, e = endoderme, c = cortex, e = épiderme et v = vacuole.

l'endoréplication sont contrôlées par des facteurs plantes présents chez certaines légumineuses (comme *M. truncatula*) ou absents chez d'autres (comme *L. japonicus*), et notamment à travers des peptides riches en cystéine (NCR) (Montiel et al., 2017). Il apparaît donc que le bactéroïde devient très dépendant de la cellule végétale pour de nombreux aspects du métabolisme cellulaire et que le maintien en tant que bactéroïde reflète un contrôle qui garantit à la plante d'héberger des bactéries qui resteront bénéfiques pour elle.

3.3. Importance du dialogue moléculaire rhizobium-légumineuse

Pour que l'interaction symbiotique soit un succès, il est nécessaire que les deux partenaires puissent se reconnaître et cette reconnaissance passe par un dialogue moléculaire complexe. Différentes molécules rentrent en jeu et sont des déterminants majeurs pour la spécificité d'hôte : les flavonoïdes et les FNs (Perret et al., 2000). La sécrétion de flavonoïdes par la plante permet l'attraction des bactéries vers la racine et va conjointement activer l'expression des gènes *nod* codant pour la synthèse des FNs requis pour la plupart des interactions rhizobium-légumineuse existantes. Dans les paragraphes qui vont suivre, ces molécules vont être plus profondément décrites pour comprendre la compatibilité entre différentes espèces de rhizobium et leur légumineuse hôte associée.

3.3.1. Les flavonoïdes

Les flavonoïdes[27] sont des métabolites spécifiques faisant partie de la grande famille des polyphénols et sont largement retrouvés dans le règne végétal. Il est estimé aujourd'hui qu'il existe plus de 10 000 composés différents appartenant à la classe des flavonoïdes, elle-même divisée en six sous-classes que sont les isoflavones, les flavones, les flavanols, les flavonols, les flavanones et les anthocyanines (Tohge et al., 2017) (figure 23). Les flavonoïdes partagent tous la même structure de base formée d'un squelette à 15 atomes de carbones (C_6-C_3-C_6) avec deux cycles aromatiques et une chaîne oxygénée. De par leur grande diversité, leurs rôles dans la rhizosphère sont multiples (Weston & Mathesius, 2013).

Au sein de la racine, les flavonoïdes s'accumulent de manière cellule-spécifique avant d'être sécrétés dans la rhizosphère qui comprend la surface de la racine mais aussi l'intérieur du cordon d'infection qui est un compartiment extracellulaire (figure 24). Ces flavonoïdes sont les premiers signaux reconnus par les rhizobiums. Par exemple la génistéine, qui a été montrée importante dans la symbiose *Bradyhrizobium-Glycine max* (Kosslak et al., 1987) ou la lutéoline et la chalcone dans le cas de

[27] Aussi appelés bioflavonoïdes, du latin "*flavus*" signifiant jaune et du grec "*eidos*" signifiant aspect. Nom donné initialement pour désigner les couleurs de la couche externe des oranges et citrons car ces molécules sont pigmentées.

l'interaction *Sinorhizobium-Medicago* (Peters et al., 1986; Liu & Murray, 2016), sont des flavonoïdes connus pour leur rôle dans la symbiose.

Parallèlement, de nombreuses études ont pu montrer que les flavonoïdes de type phytoalexines sont aussi des déterminants clés pour la spécificité d'hôte (Liu & Murray, 2016). Ces types de flavonoïdes sont connus pour avoir des activités antimicrobiennes et voient leur production augmentée durant la nodulation. Cette observation peut paraître contre-intuitive mais serait en faveur d'une sélection pour le bon rhizobium associé à sa plante hôte et jouerait sur la régulation du *quorum sensing* (Hassan & Mathesius, 2012). Enfin, des recherches menées sur le soja et *M. truncatula* ont permis de mettre en avant une catégorie de flavonoïdes[28] requis pour l'initiation et la progression de l'infection. Ils sont produits localement au niveau des sites d'infection et dans le primordium nodulaire (Liu & Murray, 2016).

NodD : un régulateur central de la nodulation

L'induction des gènes de nodulation passe par la reconnaissance et la liaison des flavonoïdes à un facteur de transcription de la famille des LysR[1] : NodD. Tous les flavonoïdes existants n'agissent pas en tant qu'inducteur des gènes *nod* et cette spécificité passe par la capacité à lier ou non NodD (Spaink et al., 1987; Li et al., 2008). La liaison du flavonoïde approprié à NodD favoriserait ainsi l'accès de l'ARN polymérase et améliorerait sa capacité de transcription des gènes *nod* sur un site spécifique du promoteur où NodD est localisé : la *nod*-box. Il a également été montré que NodD régule l'expression des gènes codant pour des systèmes de sécrétions de la bactérie (Deakin & Broughton, 2009) qui jouent des rôles contrastés dans le processus de nodulation (cf. partie 6).

L'importance du facteur de transcription NodD dans la nodulation a pu être démontrée par un mutant chez *Rhizobium leguminosarum* qui perd sa capacité symbiotique lorsque NodD n'est plus fonctionnel (Rossen et al., 1985). D'autres régulateurs ont néanmoins été découverts qui tout comme NodD peuvent moduler l'expression des gènes *nod*, positivement ou négativement (NolA, NodV et NodW) (Loh & Stacey, 2003). Enfin le fait de posséder plusieurs copies de NodD permettrait une meilleure modulation du spectre d'hôtes de la bactérie comme le suggère l'étude faite sur le transfert de NodD1 de la souche *Sinorhizobium fredii* NGR234 à d'autres rhizobiums (Bender, 1988).

[1] Ces facteurs de transcription possèdent un domaine C-terminal de fixation à un inducteur (ici les flavonoïdes) et un domaine N-terminal permettant l'interaction avec l'ADN *via* un motif hélice-coude-hélice (en anglais connu sous le terme HTH pour "helix-turn-helix").

3.3.2. Les facteurs Nod

En réponse à des flavonoïdes compatibles sécrétés par le système racinaire des légumineuses, les rhizobiums vont produire de manière souche-spécifique des molécules de type lipo-chito-oligosaccharides (LCO), aussi appelées FNs (Dénarié et al., 1996). La capacité de la plante hôte à reconnaître les "bons" FNs au cours du dialogue moléculaire entre les deux partenaires est aussi cruciale pour déterminer le spectre d'hôte. Ces molécules sont synthétisées suite à l'activation du fac-

[28] Appelés flavonoïdes d'infection.

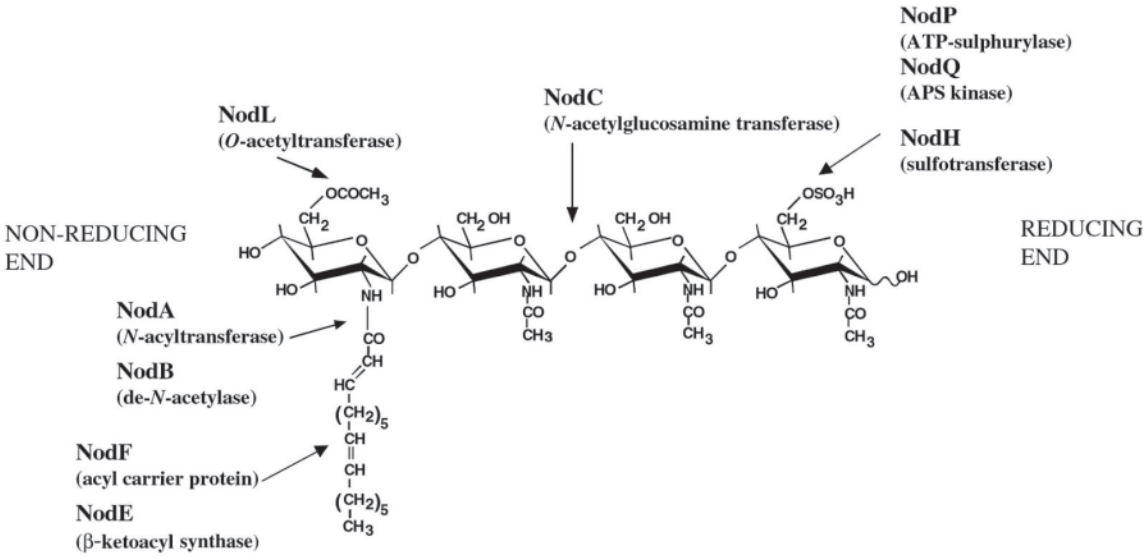


Figure 25. Structure chimique des facteurs Nod identifiés chez *S. meliloti* (Wais et al, 2002).
La structure des facteurs Nod varient d'une souche de rhizobium à une autre en fonction de ses gènes *nod* dits de « spécificités » qui vont greffer différentes décorations à une molécule de base synthétisée par les gènes *nodABC* dits communs ou canoniques. Le nombre de résidus de N-acétyl-D-glucosamine peut varier de 3 à 6 (ici 4 résidus composent le facteur Nod de *S. meliloti*), ainsi que la longueur et le niveau de saturation de la chaîne lipidique sur la partie non-réductrice de la molécule.

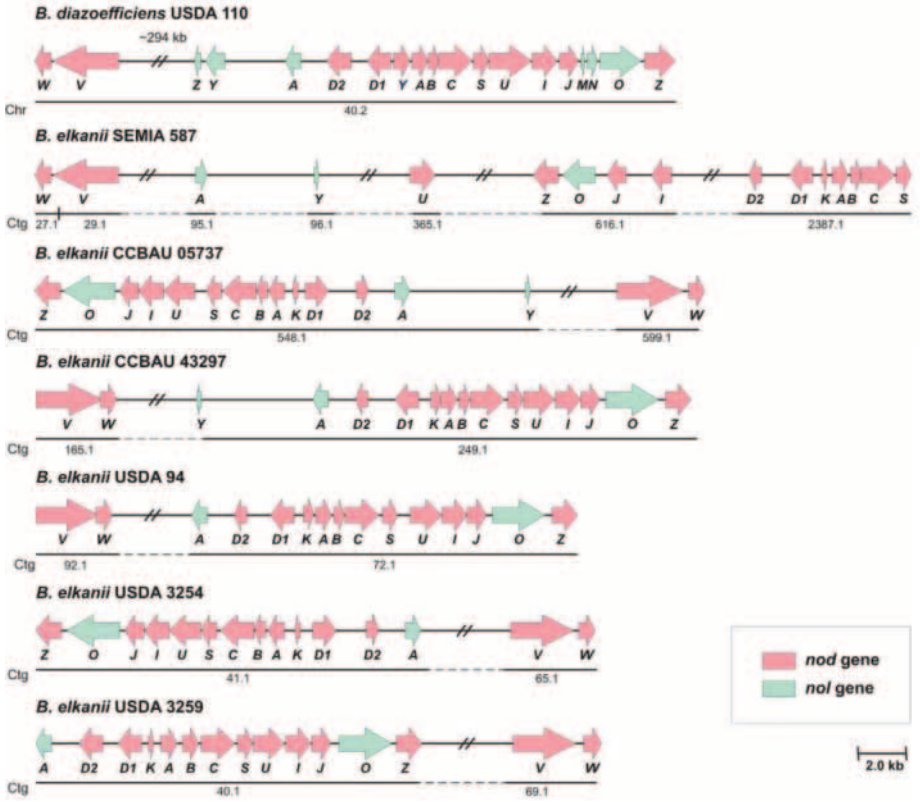


Figure 26. Organisation génétique des gènes de nodulation (*nod*) chez diverses souches de *Bradyrhizobium* (Passaglia, 2017).
Dans tous les cas, on retrouve les gènes *nodABC* canoniques entourés d'autres gènes dispensables qui participent à la spécificité des facteurs Nod de chaque souche de rhizobium donnée. Une double nomenclature *nod* et *nol* est utilisée ici pour nommer la diversité des gènes existants. On peut noter que la distance relative entre les gènes est aussi variable selon les souches.

teur de transcription NodD et sa fixation sur une *nod* box située en amont des gènes de nodulation. Les FNs ainsi produits sont ensuite sécrétés par la bactérie pour être reconnus par la plante et initier le processus de nodulation.

Les FNs sont constitués d'un squelette de trois à six résidus de N-acétyl-D-glucosamine reliés par des liaisons β-1,4[29] et comprenant une chaîne lipidique de longueur variable avec différents niveaux de saturations (Wais et al., 2002) (figure 25). Ces molécules sont synthétisées à partir de la transcription des gènes canoniques *nodABC* et seule une très faible concentration[30] est requise pour induire la nodulation. Ces gènes sont dits canoniques car ils sont généralement retrouvés chez l'ensemble des rhizobiums. Parmi les gènes canoniques, on retrouve *nodD* mais aussi les gènes *nodIJ* impliqués dans la sécrétion des FNs. D'autres gènes *nod* que ceux requis pour la synthèse des FNs contrôlent l'ajout de modifications moléculaires sur le squelette de base jouant ainsi sur la spécificité des molécules produites par une souche donnée. Ces gènes *nod* dits de "spécificités" sont multiples. Ils sont souvent organisés en opérons et décrits chez une variété de rhizobiums (Passaglia, 2017) (figure 26), certains étant communs et d'autres plus spécifiques. La nomenclature *nod* ne suffisant pas pour les décrire, ils ont été aussi nommés *noe* et *nol*. En 2009, on comptait 53 gènes de nodulations différents mais ce nombre a certainement augmenté puisque d'autres gènes ont été découverts depuis (Daubech et al., 2019) mais aucune revue récente n'a fait l'inventaire exhaustif de ces gènes à ce jour.

Trois types de modifications des FNs apportées par les gènes *nod* de spécificités existent (Mergaert et al., 1997) : (i) variation dans le degré de polymérisation du chito-oligosaccharide, (ii) variation dans la nature de l'acide gras et (iii) modifications localisées sur l'une ou les deux extrémités de l'oligosaccharide. Le spectre d'hôte est conditionné par ces différentes modifications (Lerouge et al., 1990). Par exemple, les gènes *nodH* et *nodPQ* impliqués dans la sulfatation des FNs chez *S. meliloti* sont requis pour la capacité de nodulation (Keating et al., 2002). De nombreux exemples soulignant l'importance de ces modifications dans la spécificité d'interactions ont pu être observés chez *S. meliloti*, *R. leguminosarum* et d'autres rhizobiums qui sont résumés dans la revue de Spaink (Spaink, 2000).

Par ailleurs, les rhizobiums produisent généralement un mélange de FNs de structures légèrement variables et dont la composition diffère en fonction des conditions environnementales (Mbengue et al., 2020). C'est le cas de la souche *Rhizobium* sp. NGR234 qui possède une grande diversité de structures de ses FNs et qui pourrait expliquer en partie son large spectre d'hôte (Price et al., 1992). Mais cette corrélation n'est pas toujours vraie puisque des souches possédant aussi une diversité de FNs

[29] Aussi appelée liaison osidique, la liaison β-1,4 est une des deux conformations existantes avec la liaison α-1,2 pour lier les oses (monomère de glucide) entre eux.
[30] De l'ordre du nano-molaire au pico-molaire.

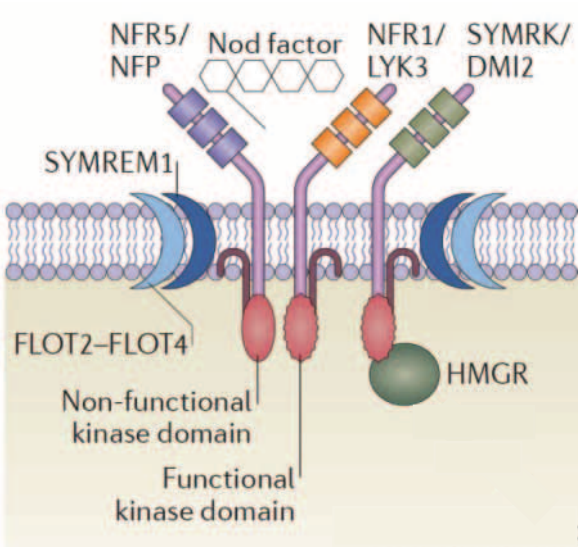


Figure 27. Un complexe de récepteurs des facteurs Nod (Oldroyd, 2013).
Représentation schématique des récepteurs identifiés chez *L. japonicus* et *M. truncatula* (les noms des protéines sont d'abord indiqués pour *L. japonicus*, suivis de ceux de *M. truncatula*). La reconnaissance du facteur Nod met en jeu un complexe de trois récepteurs NFR1/LYK3, NFR5/NFP et SYMRK/DMI2. Seuls les domaines cytoplasmiques de NFR1/LYK3 et SYMRK/DMI2 ont une activité kinase fonctionnelle. La rémorine symbiotique 1 (SYMREM1) peut s'associer avec l'ensemble de ces récepteurs et contrôle leur densité au niveau de la membrane cytoplasmique ; tout comme les flotillins FLOT2 et FLOT4 qui devraient s'associer au moins avec NFR1/LYK3. SYMRK/DMI2 peut s'associer avec HMGR (3-hydroxy-3-methylglutaryl-CoA reductase) qui est impliquée dans la production d'un messager secondaire qui pourrait être le mévalonate.

ont pourtant un spectre d'hôte réduit (Mergaert et al., 1997). Des variations dans les gènes canoniques *nodABC* peuvent aussi résulter en une spécificité d'hôte plus étroite (Roche et al., 1996).

3.4. Cascade de signalisation

La reconnaissance des FNs par la plante est une étape clé qui va permettre de déclencher toute une cascade de signalisation conduisant à l'organogenèse du nodule et à la fixation de l'azote par les bactéroïdes. Cette réponse aux FNs passe par différentes étapes telles que la dépolarisation de la membrane plasmique, une alcalinisation du cytoplasme, des flux de calcium et la déformation des poils racinaires. Les gènes impliqués dans cette réponse ont été classés dans une voie de signalisation symbiotique linéaire. Nous allons détailler dans les parties qui vont suivre les acteurs connus à ce jour de cette signalisation et leur rôle dans la mise en place de la symbiose. Cette description permettra de donner une vue générale de la voie de transduction dépendante du signal initialement orchestré par les FNs.

3.4.1. Un complexe de récepteur des FNs

La faible concentration de FNs sécrétés par la bactérie nécessite un système de récepteurs à haute affinité pour percevoir ce signal moléculaire. Il a d'abord été identifié deux récepteurs capables de lier spécifiquement les FNs : MtNFP[31] chez *M. truncatula* et LjNFR5[32] chez *L. japonicus* (Radutoiu et al., 2003; Rey et al., 2013). Ces récepteurs codent pour des sérine/thréonine kinases à motifs lysine extracellulaires[33]. Les motifs LysM sont des domaines protéiques retrouvés chez une large diversité de protéines et de récepteurs extracellulaires. Ils sont capables de lier des polysaccharides et ainsi de reconnaître tout type de molécules composées d'oses tels que les FNs. Le domaine LysM est composé d'un nombre variant de 44 à 65 acides aminés et dont la structure secondaire est associée à une paire de feuillets beta antiparallèle séparée par une paire d'hélices alpha.

Les récepteurs NFP et NFR5 possèdent trois motifs LysM qui font partie du domaine extracellulaire des protéines et leur domaine kinase intracellulaire n'est pas actif (Arrighi et al., 2006; Madsen et al., 2011). Cela suggérait donc l'existence d'autres récepteurs avec un domaine kinase actif pouvant initier la transduction du signal en interagissant avec eux. Il a alors été découvert dans ce sens un autre récepteur LysM-RLK, MtLYK3[34] et LjNFR1, qui interagit ou co-localise avec NFP et NFR5 respectivement (Smit et al., 2007; Broghammer et al., 2012). Les couples NFP/LYK3 et NFR5/NFR1 forment donc un complexe hétérodimère (figure 27). La reconnaissance des FNs déclenche l'autophosphorylation

[31] Nod Factor Perception.
[32] Nod Factor Receptor 5.
[33] Aussi abrégés récepteurs LysM-RLK pour Lysine Motif-Receptor Like Kinase.
[34] LysM-domain-containing receptor-like Kinase 3.

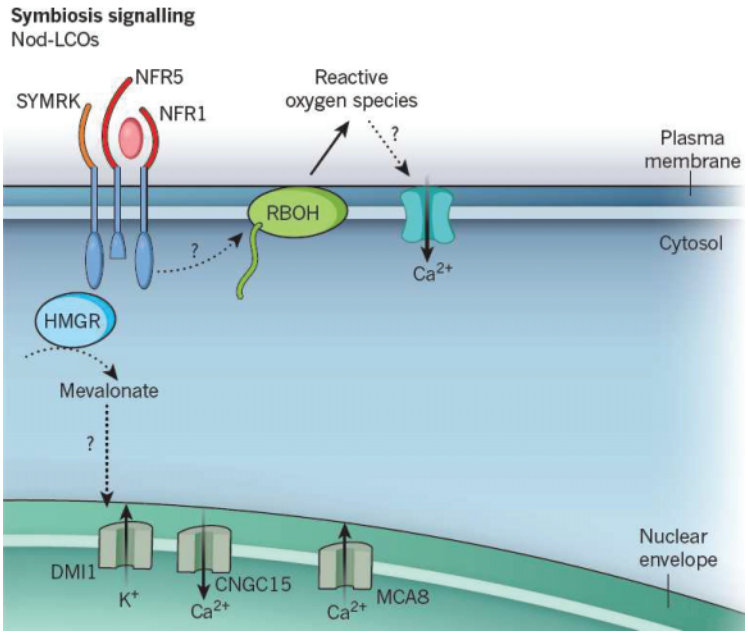


Figure 28. La perception des facteurs Nod déclenche des oscillations calciques (Zipfel et Oldroyd, 2017).
L'interaction des récepteurs suite à la reconnaissance des facteurs Nod active l'enzyme RBOH d'une manière encore inconnue. RBOH est impliquée dans la production de ROS qui active par la suite l'influx de calcium dans le cytoplasme de la cellule hôte lors d'une réponse immunitaire. En parallèle, la production d'un messager secondaire (le mévalonate) par HMGR activerait d'autres canaux calciques cette fois-ci localisés sur la membrane nucléaire. L'ensemble de ces influx-efflux de calcium provoque une série d'oscillations calciques dans le cytoplasme et le noyau qui, après décodage par CCaMK, déclenche le programme de nodulation.

de NFR1 qui transfère ensuite son groupement phosphate à NFR5 (Madsen et al., 2011). De plus, lorsqu'on exprime NFR5 et NFR1 (récepteurs issus du lotier) chez des plantes transformées de *M. truncatula*, elles sont capables d'induire des nodules en réponse à l'inoculation de *Mesorhizobium loti* qui est le symbiote normalement associé à *L. japonicus* (Radutoiu et al., 2007). Cela montre que les récepteurs des FNs jouent un rôle clé dans la spécificité d'hôte de la symbiose rhizobium-légumineuse.

Juste après la reconnaissance des FNs par leur couple de récepteur LysM-RLK, un autre récepteur de la famille des LRR[35]-RLK entre en jeu : MtDMI2[36]/LjSYMRK[37]. A l'inverse du mutant *nfp*, la mutation de *dmi2* présente toujours une réponse aux rhizobiums et aux FNs mais aucun nodule n'est observé, suggérant qu'il agirait en aval du complexe NFP/LYK3 (Stracke et al., 2002). En 2014, il a pu être montré que SYMRK interagit au niveau de la membrane plasmique avec NFR5 (Antolín-Llovera et al., 2014) et que son domaine kinase reconnaît une enzyme[38] clé de la voie de biosynthèse du mévalonate (figure 27). Ce métabolite participerait à la transduction du signal au sein de la cellule du poil absorbant jusqu'au noyau pour produire une série d'oscillations calciques qui est alors décryptée pour activer les gènes de nodulation. En effet, son application seule permet d'activer un signal calcique chez la plante et serait donc un messager secondaire après la reconnaissance des FNs (Venkateshwaran et al., 2015; Balloon et al., 2015). D'autres partenaires de DMI2/SYMRK ont été identifiés et joueraient un rôle dans leur renouvellement et leur dégradation telle que l'E3 ubiquitine ligase PUB1[39], capable d'interagir aussi avec LYK3 (Vernié et al., 2016).

La mobilité de ces récepteurs sur la membrane plasmique est assurée par des radeaux lipidiques tels que FLOT2 et FLOT4 (Haney & Long, 2010). Des rémorines[40] sont également associées à ces radeaux lipidiques et SYMREM1[41] apparaît important pour l'infection rhizobienne (figure 27). La capacité de SYMREM1 à s'associer avec LYK3 (Lefebvre et al., 2010) suggère que ces acteurs sont importants pour faciliter la formation de complexes de récepteurs et contrôler l'intensité de la signalisation afin d'adapter la réponse aux FNs (Oldroyd, 2013).

En parallèle, des ROP-GTPases[42] interagissent avec les récepteurs LysM-RLK durant l'infection par les rhizobiums (Ke et al., 2012). Cette interaction déclenche l'activation d'une RBOH[43] qui est une enzyme connue pour produire ROS et activer l'influx de calcium dans la cellule lors d'une réponse im-

[35] Leucine-rich repeats.
[36] Does not Make Infections 1, aussi connu sous le nom de NORK (Nodulation Receptor Kinase).
[37] Symbiosis Receptor-like Kinase.
[38] HMGR qui est une 3-hydroxy-3-methylglutaryl-CoA reductase.
[39] Plant U-box.
[40] Protéines spécifiques du règne végétal connues pour être localisées à la membrane plasmique.
[41] Symbiotic Remorin 1.
[42] Rho-related GTPases from plants, protéines de la famille Rho capable de lier le GTP (guanosine triphosphate) et de l'hydrolyser en GDP (guanosine diphosphate).
[43] RBOH est une NADPH oxidase.

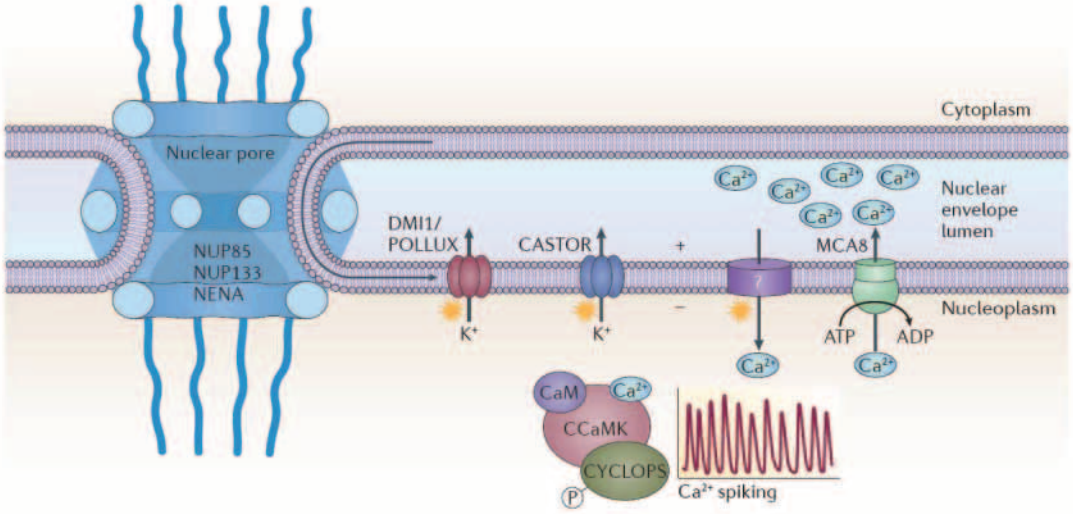


Figure 29. Une machinerie complexe responsable des oscillations calciques nucléaires (Oldroyd, 2013).
En aval des récepteurs symbiotiques, diverses protéines sont responsables des oscillations calciques et sont associées à la membrane nucléaire. MtDMI1/LjPOLLUX et CASTOR seraient préférentiellement localisés sur la membrane interne du noyau. Cette localisation serait facilitée par le pore nucléaire nécessitant trois composants NUP85/NUP133/NENA qui sont nécessaires à la génération du signal calcique. Un autre canal qui reste à définir permettrait de maintenir un équilibre de la charge membranaire *via* une ATPase (MCA8) qui utilise de l'ATP pour ramener le calcium dans l'enveloppe nucléaire. L'étoile jaune indique une activation possible du canal par un ligand. Par ailleurs, il a été montré plus récemment que DMI1/POLLUX serait finalement un transporteur de calcium et non de potassium (S. Kim et al., 2019). L'ensemble de cette machinerie induit des oscillations nucléaires qui sont déchiffrées par MtDMI3/LjCCaMK pour activer par la suite une activation transcriptionnelle des gènes de nodulation. Cette activation passe entre-autre par l'interaction de CYCLOPS avec CCaMK.

munitaire de la plante. Dans le cas de la symbiose, RBOH promeut l'infection et la nodulation (Marino et al., 2011; Montiel et al., 2012). Cette production de ROS activerait aussi un canal calcique inconnu localisé sur la membrane plasmique pour la rentrée de Ca^{2+} dans le cytoplasme (figure 28). De manière intéressante, cet influx calcique nécessite des concentrations plus élevées en FNs que pour les oscillations calciques nucléaires (Morieri et al., 2013). Ce processus est donc mis en place plus tardivement, pendant l'infection par les rhizobiums et après leur piégeage dans le poil absorbant où ils sont plus résistants au stress oxydatif produit par les ROS (Zipfel & Oldroyd, 2017). Les quantités de FNs sécrétés par les rhizobiums doivent ainsi être finement régulées pour que la bactérie puisse pénétrer les tissus de la plante.

Il a été récemment montré chez le lotier que l'interaction de SYMRK avec BAK1[44], un régulateur positif de l'immunité de la plante, permet d'inhiber l'activité kinase de BAK1. Cette interaction entre les deux récepteurs est promue par l'inoculation de la bactérie qui diminue ainsi la réponse immune de la plante (Feng et al., 2021). Ce lien étroit entre immunité et symbiose sera discuté dans la partie 4 de cette introduction.

3.4.2. Déclenchement et décryptage du signal calcique

La surexpression de SYMRK ou du domaine kinase de DMI2 conduit à la formation de nodules en absence de la bactérie (Ried et al., 2014). Néanmoins, ce phénotype est dépendant des gènes *MtDMI3*/*LjCCaMK*[45] et *MtDMI1*/*LjCASTOR-LjPOLLUX*. En effet, la transduction du signal par SYMRK en réponse aux FNs est médiée par des oscillations calciques produites au niveau de la membrane nucléaire par les canaux calciques CNGC15a, CNGC15b et CNGC15c[46] chez *M. truncatula* (Charpentier et al., 2016) ou NUP85, NUP133[47] et NENA chez *L. japonicus* (Kanamori et al., 2006; Saito et al., 2007; Groth et al., 2010). Ces canaux calciques sont capables d'interagir avec DMI1/CASTOR et POLLUX qui sont eux aussi des transporteurs de calcium bien que DMI1 ait été initialement proposé comme transporteur de potassium (Kim et al., 2019). Cette interaction permet de contrôler finement l'influx et l'efflux de Ca^{2+} dans le noyau (figure 29). Une ATPase nucléaire, MCA8[48] chez *M. truncatula*, maintient quant à elle la concentration en calcium nucléaire (Capoen et al., 2011). Cette repolarisation de la membrane nucléaire induite par ces flux ioniques déclenche les oscillations calciques. Ainsi, le signal de nodulation initialement cytoplasmique est transmis au noyau. Ces séries d'oscillations calciques sont alors déchiffrées par DMI3/CCaMK au sein du noyau pour initier une réponse transcrip-

[44] Brassinosteroid insensitive 1-Associated receptor Kinase 1.
[45] Calcium/CalModulin-dependent protein Kinase.
[46] Cyclic Nucleotide-Gated Channels a/b/c.
[47] Nucleoporin 85/133.
[48] Medicago Ca^{2+} ATPase.

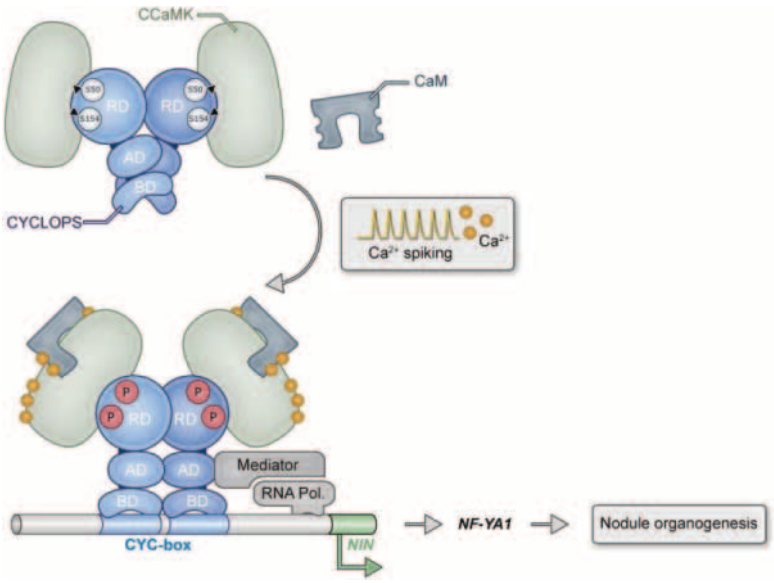


Figure 30. Modèle proposé pour le fonctionnement du complexe CCaMK/CYCLOPS dans le décodage des oscillations calciques conduisant à l'organogénèse nodulaire (Singh et al, 2014).
CCaMK et CYCLOPS forment un complexe inactif dans le noyau où CYCLOPS est présent sous forme de dimère. En réponse au rhizobium, le signal calcique active CCaMK *via* la calmoduline (CaM) et la liaison avec le Ca2+. CCaMK phosphoryle ensuite CYCLOPS sur des sérines spécifiques, ce qui induit un changement de conformation et libère le domaine de liaison à l'ADN (BD). Dès lors, CYCLOPS peut se lier à l'ADN au niveau d'une CYC-box située dans le promoteur de NIN. Cette liaison à l'ADN implique le recrutement de la machinerie de transcription (complexe médiator et ARN polymérase II) *via* son domaine AD. L'expression de NIN puis de NF-YA1 conduit à l'organogénèse nodulaire.

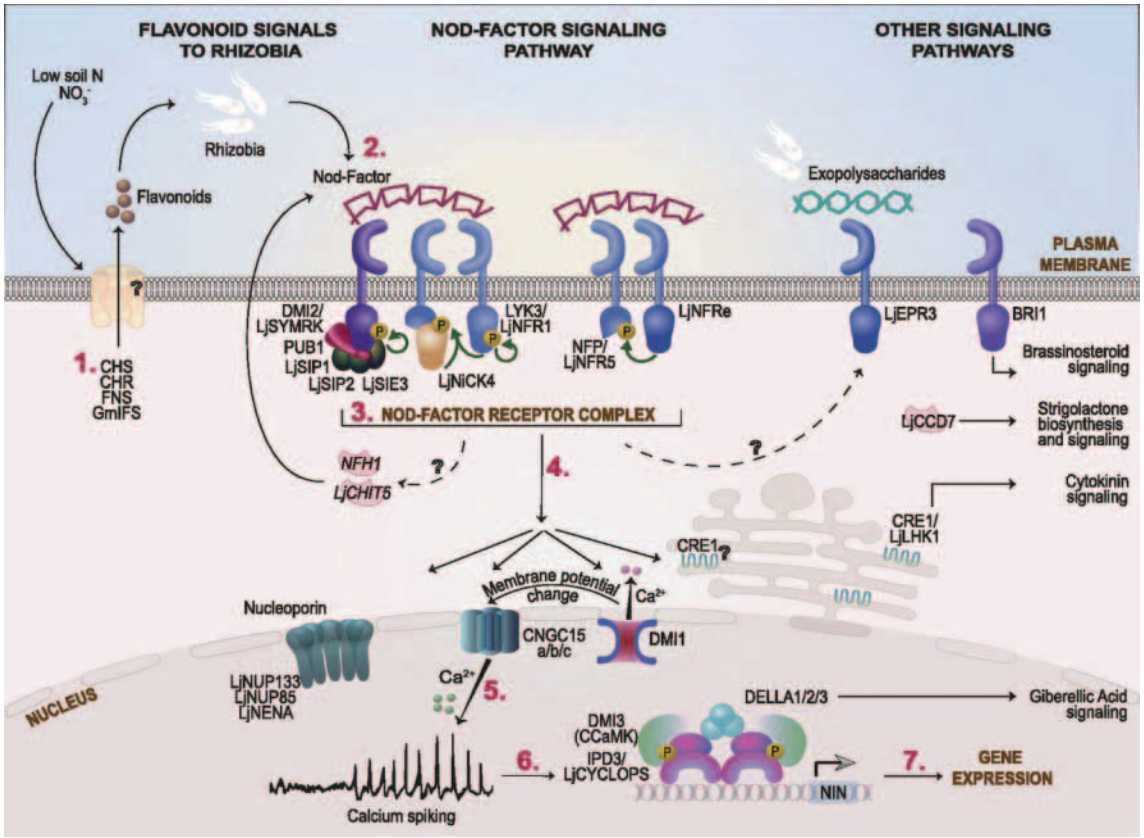


Figure 31. Gènes et processus impliqués dans la signalisation précoce au cours de la nodulation (Roy et al, 2020).
La production de flavonoïdes (1) par la plante active la synthèse des facteurs Nod (2) chez la bactérie qui vont ensuite être reconnus par un complexe de récepteurs (3). S'ensuit alors un signal calcique médié par divers acteurs (4-6) qui va conduire à l'expression des gènes de nodulation (7).

tionnelle et cette signature calcique doit avoir une durée et une fréquence spécifique pour induire le processus développemental (Charpentier, 2018; Leitão et al., 2019).

L'activation de CCaMK dans le noyau déclenche le programme symbiotique en entier, incluant l'organogenèse nodulaire (Lévy et al., 2004). Cette activation passe par la capacité de CCaMK à lier le calcium sous deux formes : i) une forme libre du Ca^{2+} se liant directement au domaine EF-hand[49] et ii) une forme complexée du Ca^{2+} qui fixe la calmoduline (CaM) (Gleason et al., 2006). S'ensuit alors une autophosphorylation de CCaMK par son domaine sérine/thréonine kinase.

Ainsi, la perception des FNs conduit dans les minutes qui suivent à la production dans le noyau d'oscillations calciques qui sont caractéristiques de la transduction du signal symbiotique. CCaMK est le régulateur central du développement symbiotique de la racine en traduisant ce signal (figure 29). En effet, les mutants *ccamk* ou *dmi3* sont affectés à la fois dans l'infection et la nodulation bien que les oscillations calciques soient toujours présentes (Mitra et al., 2004). De plus, des formes auto-actives de CCaMK ou DMI3 sont capables d'induire des nodules spontanés en l'absence de la bactérie et de restaurer la nodulation chez des mutants affectés dans le signal calcique soulignant l'importance de cet acteur moléculaire dans la symbiose (Gleason et al., 2006; Tirichine et al., 2006).

3.4.3. Expression des premiers gènes de nodulation

La recherche des cibles de CCaMK a permis d'identifier chez les légumineuse modèles *MtIPD3*[50]/*LjCYCLOPS* (Messinese et al., 2007; Yano et al., 2008). Ces gènes codent pour des activateurs de la transcription et se trouvent être activés par CCaMK par transfert du groupement phosphate issu de son autophosphorylation (Yano et al., 2008). Il a pu être montré que CYCLOPS se lie à l'ADN de manière séquence-spécifique et dépendant de la phosphorylation de deux sérines pour activer la transcription de gènes cibles (Singh et al., 2014). Il active ainsi le facteur de transcription NIN qui lui-même active le complexe NF-Y[51] pour initier le processus symbiotique nodulaire (Soyano et al., 2013) (figure 30). Par ailleurs, la forme phosphorylée de CYCLOPS est suffisante pour déclencher l'organogenèse nodulaire en absence de la bactérie et indépendamment de CCaMK (Singh et al., 2014). Parallèlement, le mutant *cyclops* montre un niveau d'expression de NIN plus faible et devient incapable de noduler. Cependant le phénotype de ce mutant n'est pas tout à fait le même que celui du mutant *ccamk*, soulignant l'existence d'autres acteurs à ce niveau de la cascade de signalisation symbiotique. Il est ainsi possible qu'il existe des cibles alternatives de CYCLOPS proches de NIN qui sont requises pour la formation du nodule (Limpens & Bisseling, 2014).

[49] Motif largement retrouvé chez les protéines capables de lier et de réguler le calcium. Sa structure hélice-boucle-hélice permet de fixer un ion calcium et peut rappeler la forme d'une main.
[50] Interacting Protein of DMI3.
[51] Nuclear Factor-Y, complexe protéique très répandu chez les eucaryotes.

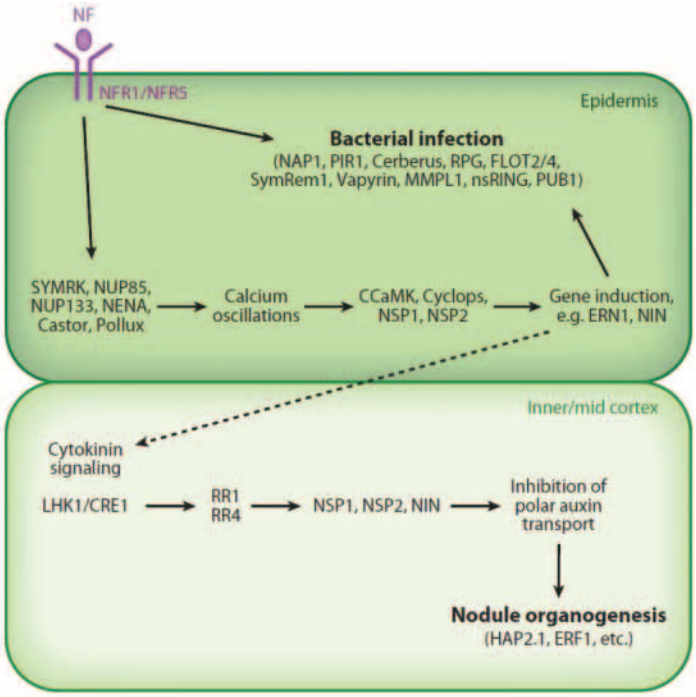


Figure 32. Représentation schématique des voies génétiques impliquées dans l'organogénèse des nodules et l'infection bactérienne (Oldroyd, 2011).
La perception des facteurs Nod a lieu au niveau de l'épiderme de la racine ce qui va déclencher toute une voie de signalisation jusqu'à l'expression de facteurs de transcription tels que NIN et ERN1 pour activer l'expression de gènes spécifiques. Cette voie de signalisation va conjointement initier l'infection bactérienne au niveau de l'épiderme à travers l'expression de gènes requis pour l'infection, ainsi qu'activer les divisions cellulaires dans le cortex par l'intermédiaire d'un signal diffusible inconnu (ligne en pointillées).

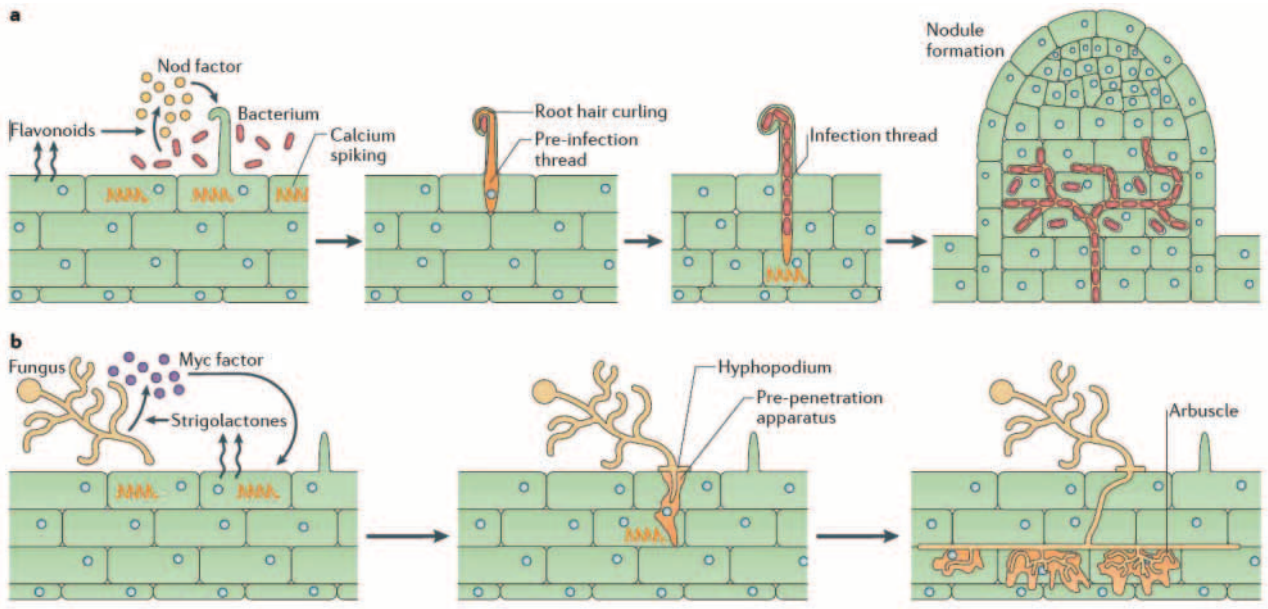


Figure 33. Comparaison des première étapes de colonisation racinaire entre un rhizobium et un champignon mycorhizien à arbuscule jusqu'à la mise en place de nouvelles structures pour échanger des nutriments avec la plante (Oldroyd, 2013).
(**a**) Schéma du processus classique de nodulation entre une légumineuse et un rhizobium. (**b**) Schéma du processus de mycorhization à arbuscules entre une plante et un champignon. Dans les deux cas, des molécules sont sécrétées par la plante (flavonoïdes/strigolactones) qui sont reconnus par le partenaire symbiotique et qui conduit à la production de molécules signalétiques (facteurs Nod/facteurs Myc) qui seront reconnues par la plante à travers des récepteurs spécifiques. La pénétration du symbiote s'initie au niveau des cellules épidermales jusqu'à former de nouvelles structures dans le cortex (nodule/arbuscule).

Pour son activité transcriptionnelle CYCLOPS n'agit pas seul mais au sein d'un complexe protéique plus large constitué de plusieurs facteurs de transcription contenant un domaine GRAS[52] (figure 31), tels que NSP1/NSP2 et les protéines DELLA qui sont aussi requis pour l'infection rhizobienne (Fonouni-Farde et al., 2016; Jin et al., 2016). Ce complexe est connu pour activer d'autres facteurs de transcriptions que NIN comme ERN1[53] qui est requis pour la formation du cordon d'infection (Cerri et al., 2017). L'activation complète de cette voie de signalisation, depuis les récepteurs LysM reconnaissant les FNs jusqu'à l'expression de facteurs de transcription tels que NIN et ERN1, conduit à deux réponses développementales différentes : épidermale ou corticale (figure 32). Ces deux réponses sont activées conjointement bien qu'elles se déroulent au niveau d'assises cellulaires différentes. Au niveau de l'épiderme a lieu l'initiation de l'infection bactérienne tandis que dans le cortex sont initiées les premières divisions cellulaires conduisant à l'organogenèse nodulaire. Il existe donc une communication coordonnée entre les cellules de l'épiderme et celles du cortex qui passe par le transport symplasmique *via* les plasmodesmes[54], d'un signal qui reste encore à identifier (Complainville et al., 2003; Gaudioso-Pedraza et al., 2018).

Au niveau de l'épiderme, l'induction des gènes de nodulation active l'expression de divers acteurs requis pour l'infection. On peut par exemple citer MtLIN[55]/LjCERBERUS (Kiss et al., 2009; Yano et al., 2009) ou la VAPYRIN (Murray et al., 2011) qui sont impliquées dans le bon développement du cordon d'infection. Au niveau du cortex, l'activation de la voie de signalisation symbiotique conduit à une signalisation médiée par les cytokinines induite par le récepteur MtCRE1/LjLHK1. Cette hormone joue un rôle clé dans l'organogenèse nodulaire en supprimant le transport polaire d'auxine. D'autres acteurs jouent un rôle lors de la réponse corticale comme par exemple les protéines MtRR1 et MtRR4[56] qui agissent respectivement de manière positive ou négative sur la nodulation en aval des cytokinines (Gonzalez-Rizzo et al., 2006; Plet et al., 2011).

3.4.4. Similarités avec la symbiose mycorhizienne : voie de signalisation symbiotique commune

Il est intéressant de remarquer qu'il existe des similarités morphologiques et génétiques entre les endosymbioses nodulaires et l'endomycorhization mettant en jeu des champignons mycorhiziens à arbuscules (Martin et al., 2017). Dans les deux cas, le symbiote pénètre les tissus de la plante, suite à

[52] Nommé ainsi suite aux trois premiers gènes caractérisés possédant ce motif : *AtGAI*, *AtRGA*, *AtSCR*. La famille de facteurs de transcription GRAS est l'une des plus représentée chez les plantes.
[53] ERF Required for Nodulation 1.
[54] Canaux traversant la paroi végétale et qui permettent le passage d'eau, de solutés, de phytohormones et aussi de virus. Son diamètre peut varier entre 20 et 40 nm. Leur ouverture est contrôlée par le cytosquelette.
[55] Lumpy Infections.
[56] Response Regulator 1 et 4. RR1 est de type B tandis que RR4 est de type A.

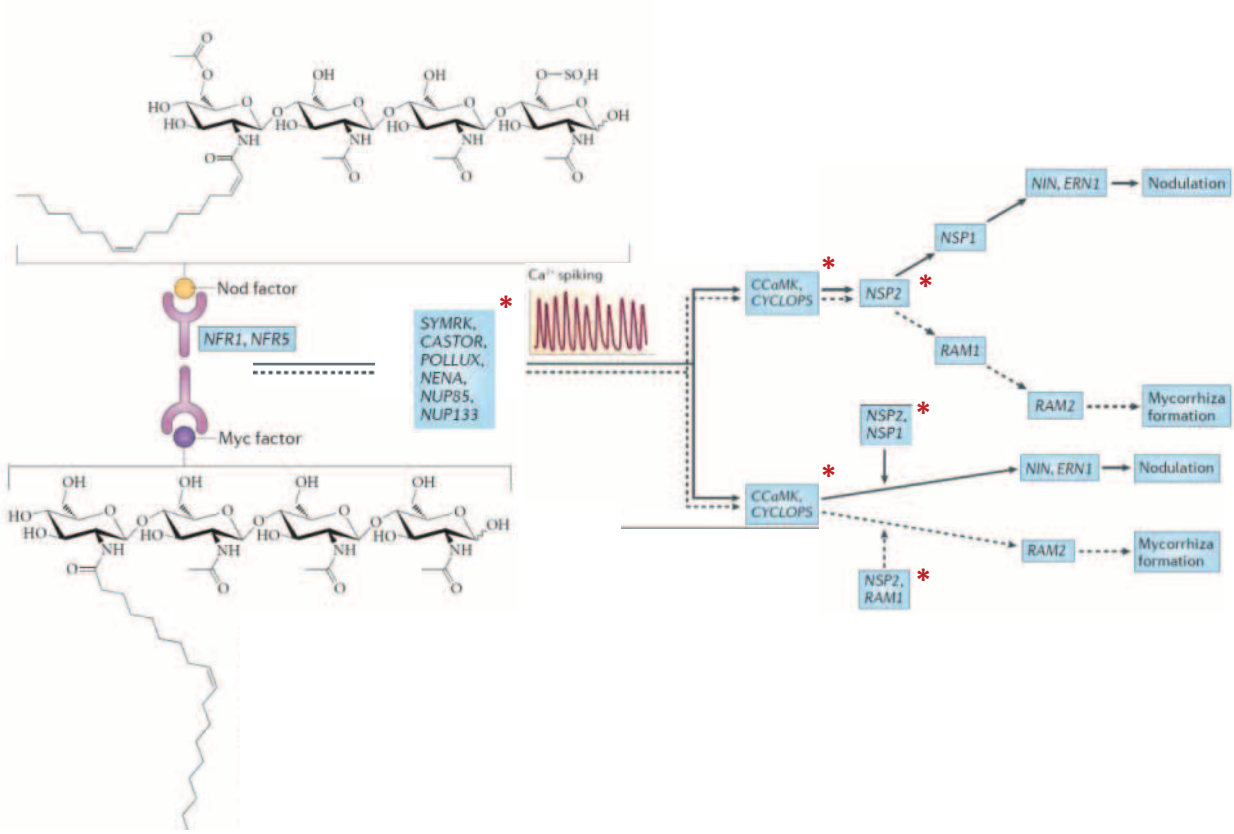


Figure 34. Voie de signalisation symbiotique partagée entre la nodulation et la mycorhization (adapté de Oldroyd, 2013). La structure des facteurs Nod (ici celui de *S. meliloti*) et celle des facteurs Myc (ici celui de *Glomus intraradices*) sont similaires. Ces molécules induisent une voie de signalisation symbiotique dont le nom des gènes sont indiqués pour *L. japonicus*. La voie empruntée par la nodulation est indiqué par des flèches pleines et celle empruntée par la mycorhization par des flèches en pointillées. L'étoile rouge indique les gènes conservés entre les deux processus symbiotiques.

la reconnaissance d'exsudats racinaires et va mettre en place des structures adaptées pour l'échange de nutriments entre les deux partenaires (Denison & Kiers, 2011; Parniske, 2008) (figure 33). Lors de la symbiose nodulaire, ce sont les flavonoïdes qui sont sécrétés par les racines et pour la symbiose mycorhizienne ce sont les strigolactones[57] qui déclenchent la signalisation symbiotique. En réponse à ces molécules, le champignon synthétise aussi des LCOs appelés facteurs Myc dont la structure est similaire à celle des FNs (figure 34). Autres points communs, ce sont les gènes de la plante requis pour la mise en place de la symbiose : au moins 7 gènes communs ont été montrés nécessaires et codent pour des protéines signalétiques impliquées dans la mise en place du nouvel organe hébergeant le symbiote[58] (Kistner et al., 2005). Ces gènes sont les suivants : SYMRK, CASTOR, POLLUX, NUP85/133/NENA, CCaMK, CYCLOPS et NSP2 (Oldroyd, 2013). On parle alors de voie de signalisation symbiotique commune[59] (figure 34). Cependant, le nombre de gènes conservés entre les deux processus symbiotiques dépasse aujourd'hui celui initialement identifié. Par exemple, les protéines DELLA mais aussi CERBERUS ou encore VAPYRIN participent aussi à ces deux symbioses (Jin et al., 2016; Murray et al., 2011; Takeda et al., 2013). Enfin l'établissement de ces deux types de symbioses permet aux plantes de faire face à des carences nutritives. Le phosphate étant l'élément majoritaire rendu biodisponible pour les plantes mycorhizées et l'azote pour les plantes nodulées.

Néanmoins, alors que les symbioses nodulaires n'impliquent qu'un groupe restreint d'hôtes, l'endomycorhization est une symbiose bien plus répandue puisqu'il est estimé que 70-90% des plantes terrestres sont mycorhizées (Parniske, 2008). Enfin, il a été admis qu'un ancêtre commun aux deux types d'endosymbioses citées est apparu il y a 450 millions d'années bien que la symbiose fixatrice d'azote serait plus récente. Elle résulterait d'un unique précurseur au sein des Rosidae (ou Rosidées) apparu il y a un peu plus de 100 millions d'années (Werner et al., 2014). Ultérieurement, d'autres évènements évolutifs ont conduit à différents types de symbioses nodulaires distincts chez les angiospermes, expliquant en partie les différences entre les modes d'infection, les formes des nodules, l'identité du symbiote et les processus de contrôle de cette symbiose (Kistner & Parniske, 2002; Werner et al., 2014). De manière assez contre-intuitive et bien qu'apparaissant comme un avantage évolutif pour les plantes, la capacité à développer une symbiose fixatrice d'azote a été plusieurs fois perdues au cours de l'évolution soulignant la fragilité de ce trait évolutif qui n'est pas forcément un avantage sélectif (Griesmann et al., 2018; Werner et al., 2014).

[57] Hormones végétales découvertes en 1960 et possédant une structure cyclique de trois lactones.

[58] Ces déterminants communs sont aussi retrouvés dans le cas des symbioses actinorhiziennes (Hocher et al., 2011; Perrine-Walker et al., 2011).

[59] En anglais, Common Symbiotic Signaling Pathway (CSSP).

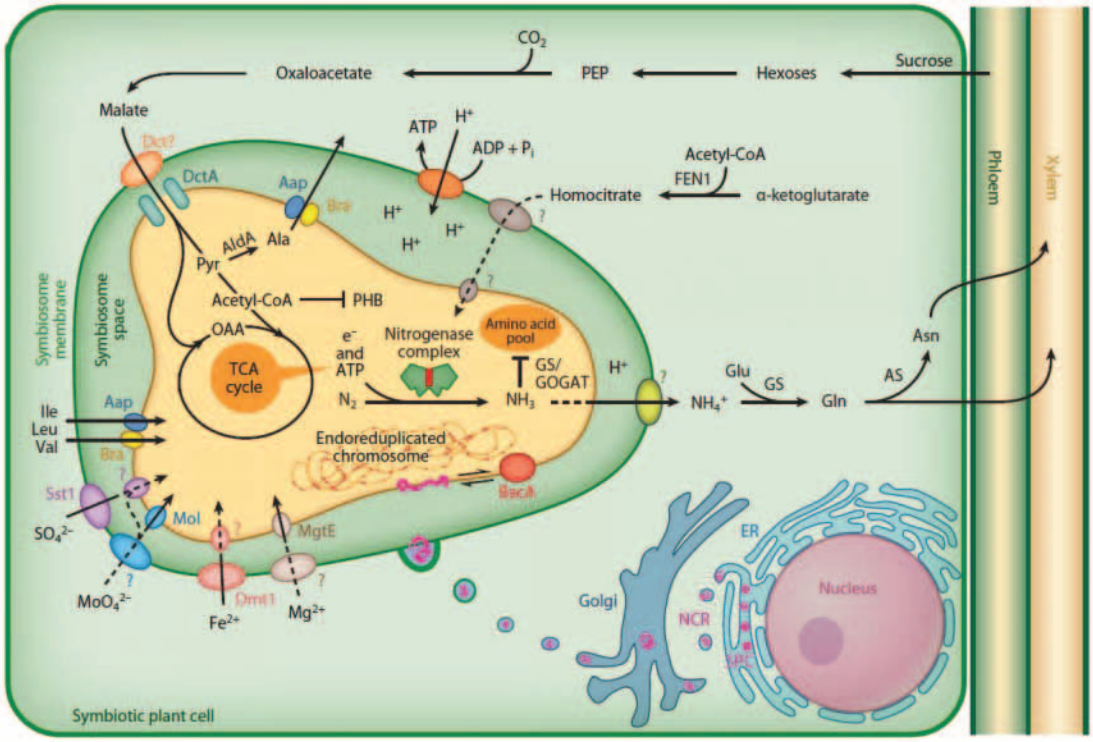


Figure 35. Fonctionnement d'un bactéroïde au sein du symbiosome (Oldroyd, 2011).
Parmi les éléments apportés par la plante pour le bon fonctionnement du bactéroïde, on observe l'apport de malade comme une des sources carbonées et qui permet le fonctionnement du cycle de Krebs (ou TCA). Cette énergie produite sous forme d'ATP assure le bon fonctionnement du complexe enzymatique de la nitrogénase qui requiert l'apport d'homocitrate fournit par la plante. En échange, le bactéroïde suspend son assimilation d'ammoniac pour le donner à la plante. Cette dernière va assimiler l'azote à travers son intégration dans les acides aminés. Enfin, pour le bon fonctionnement de la nitrogénase, l'apport d'ions métalliques tels que le Fe^{2+} ou le Mo sont requis. De nombreux transporteurs sont donc requis pour ces échanges et ceux décrits ici ont été découverts chez plusieurs espèces différentes (*Rhizobium leguminosarum*, *Lotus japonicus*, *Glycine max*, *Medicago truncatula*). Dans les cas des légumineuses appartenant à la famille des Papilionoideae, de nombreux peptides NCR sont transportés puis sécrétés dans le symbiosome à différentes étapes du développement du nodule.

3.5. Métabolisme du nodule

La dernière étape pour qu'un nodule soit pleinement fonctionnel est la mise en place d'un système d'échanges entre les bactéroïdes fixateurs d'azote et la plante donneuse de photosynthétats sous forme de carbone réduit. Des études génétiques ont montré que la symbiose fixatrice d'azote nécessite une coordination de diverses voies métaboliques et de multiples transporteurs entre la plante et le rhizobium. Nous allons décrire de manière générale dans les parties qui vont suivre les mécanismes requis pour une symbiose fixatrice d'azote efficiente à bénéfice réciproque, qui sont plus détaillés dans les revues suivantes : Oldroyd et al., 2011; Udvardi & Poole, 2013; Poole et al., 2018; Roy et al., 2020.

3.5.1. Fixation d'azote par la nitrogénase et taux d'oxygène

L'environnement cellulaire du symbiosome nécessite des conditions spécifiques pour un bon fonctionnement de la nitrogénase. Le métabolisme du nodule doit donc être contrôlé finement pour favoriser la fixation d'azote (figure 35). Il a pu être montré que des taux élevés de respiration cellulaire conduit à une production de ROS nocifs pour le nodule (Chang et al., 2009). C'est pourquoi de nombreux antioxydants tels que l'ascorbate ou l'homoglutathion sont retrouvés dans les nodules pour piéger les ROS et protéger ainsi les cellules de stress oxydatifs. Le zinc joue également un rôle dans la détoxification des ROS en faisant partie des superoxydes dismutases[60] ; en plus d'être présent chez des facteurs de transcription à doigts de zinc requis pour la nodulation (González-Guerrero et al., 2014). Ainsi, divers transporteurs de zinc localisés à la membrane plasmique des cellules infectées ont été identifiés (Abreu et al., 2017; León-Mediavilla et al., 2018).

L'apport d'autres ions métalliques tels que le fer ou le molybdène est crucial pour la fixation du diazote par la nitrogénase (Roy et al., 2020). Dans le cas du fer, il est délivré dans le nodule à travers le système vasculaire avant d'être transmis aux bactéroïdes (Rodríguez-Haas et al., 2013). Le transporteur de fer NRAMP1[61] joue un rôle clé pour la fixation symbiotique de l'azote chez *M. truncatula* par exemple (Tejada-Jiménez et al., 2015). Des transporteurs de molybdène ont aussi été identifiés (MtMOT1.2 et MTMOT1.3[62]) et sont situés au niveau de la membrane plasmique de l'endoderme entourant le nodule (Gil-Díez et al., 2019).

La nitrogénase nécessite des molécules d'ATP pour catalyser la réaction enzymatique de la réduction de l'azote. Cette énergie lui est fournie par la plante et idéalement, un fort flux d'oxygène permettrait des taux élevés de synthèse de l'ATP. Cependant, l'environnement du symbiosome doit être

[60] Métalloprotéines catalysant la dismutation des anions superoxydes en oxygène et peroxyde d'hydrogène.
[61] Resistance-Associated Macrophage Protein 1.
[62] Molybdate Transporter 1.2 et 1.3.

pauvre en oxygène afin d'éviter l'inactivation de la nitrogénase. C'est à ce moment que la léghémoglobine[63] intervient (Downie, 2005). Ce sont les protéines végétales les plus abondantes du nodule (Udvardi & Poole, 2013) et jouent un rôle crucial dans le maintien d'une pression partielle faible en oxygène par leur capacité à lier l'O_2 et à l'exporter en dehors de la zone de fixation du N_2 (Ott et al., 2005). Leur présence permet ainsi à la bactérie d'effectuer ces cycles respiratoires pour produire de l'ATP tout en maintenant un équilibre redox précis et favorable au fonctionnement de la nitrogénase. La production de léghémoglobine dans le nodule donne cette couleur rosée/rouge caractéristique d'un nodule fixateur.

3.5.2. Auxotrophie symbiotique

Les bactéroïdes en développement ou matures peuvent devenir dépendants de la plante pour l'apport de certains acides aminés comme cela a pu être observé chez le pois ou le haricot (Prell et al., 2009). En effet, les transporteurs de type ABC pour le transport d'acides aminés à chaîne ramifiée ont été montré essentiels pour une bonne fixation d'azote et leur mutation conduit à un développement anormal des bactéroïdes (Prell et al., 2009). On parle alors d'auxotrophie symbiotique car sous la forme de bactéroïde, la synthèse des acides aminés est stoppée et nécessite donc l'apport de ces composés par la plante (ce qui n'est pas le cas en mode de vie libre). L'activation de la nitrogénase conduisant à la synthèse d'ammoniac par les bactéroïdes s'accompagne d'un arrêt de l'assimilation du NH_3 en acides aminés, ce qui explique cette dépendance vis-à-vis de la plante. La mutation du système GOGAT[64] chez des bactéries du genre *Rhizobium* rend l'assimilation de l'ammoniac impossible mais n'affecte cependant pas la fixation d'azote (Mulley et al., 2011). Le rôle d'un bactéroïde au sein du symbiosome est donc de produire du NH_3 sans l'utiliser. Par ailleurs d'autres déterminants tels que l'homocitrate synthase et les NCR participent à cette auxotrophie symbiotique et ont été discutés précédemment.

3.5.3. Echanges nutritionnels entre la plante et la bactérie

Une des formes de carbone transportées par le système vasculaire de la plante jusqu'aux nodules est le saccharose. Dérivé de la photosynthèse, il est ensuite hydrolysé par des sucroses synthases ou des invertases présents en plus grand nombre dans les nodules (Flemetakis et al., 2006; Hohnjec et al., 2003). Cette étape d'hydrolyse de ce disaccharide est importante pour une symbiose efficiente. En

[63] Association du terme légumineuse et hémoglobine. Hémoprotéine fixatrice du dioxygène complexant un atome de fer qui lui confère une couleur rouge rappelant l'hémoglobine des vertébrés.

[64] Glutamine oxoglutarate aminotransferase, voie métabolique qui permet l'assimilation de l'azote NH_3/NH_4^+ sous forme organique en produisant du glutamate. Cette enzyme est étroitement associée à la glutamine synthétase (GS). On parle alors d'assimilation d'azote par la voie GS-GOGAT.

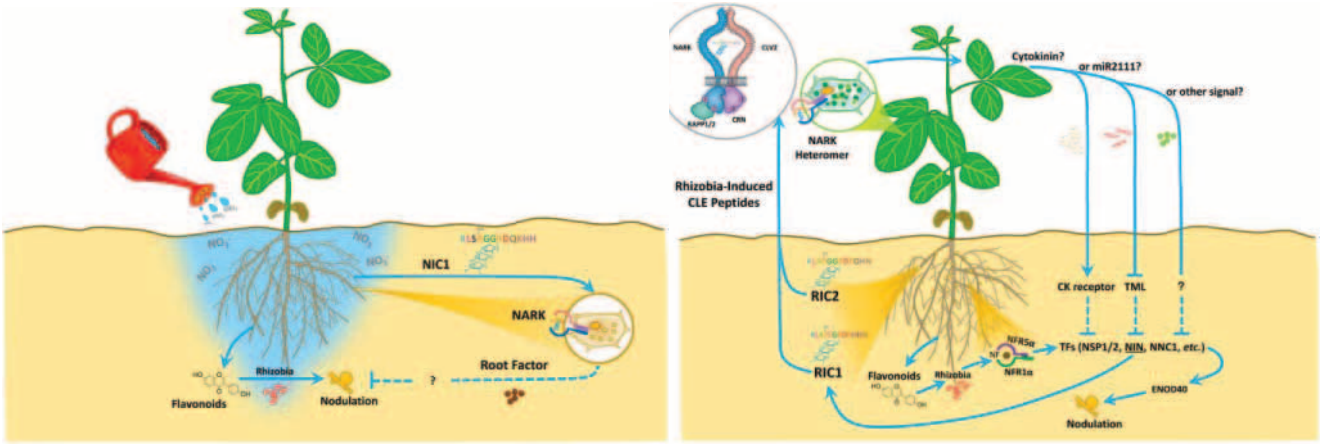


Figure 36. Mécanismes de régulation de la nodulation (Ferguson et al., 2019).
(A gauche) La présence de nitrate dans les sols régule négativement la nodulation notamment par la production de peptides CLE (par exemple NIC1 chez le soja). Ces peptides sont perçus localement par le récepteur kinase NARK (chez le soja) conduisant à la production de facteurs encore inconnus qui vont restreindre la nodulation. (A droite) L'autorégulation de la nodulation passe aussi par des peptides CLE (tels que RIC1 et RIC2) qui sont transportés dans les parties foliaires où ils sont perçu par NARK au sein d'un complexe protéique (CLAVATA2, CORYNE et/ou KLAVIER). La phosphorylation/déphosphorylation de NARK transduit le signal qui *in fine* régule négativement des acteurs requis pour la nodulation par des mécanismes encore inconnus.

effet, des mutants dans les enzymes catalysant cette réaction voient leur fixation d'azote diminuer (Baier et al., 2007). La plante fournit alors du malate aux bactéroïdes qui est issu de la transformation du saccharose en diverses molécules intermédiaires (Booth et al., 2021). Ce malate est ensuite utilisé par le cycle de Krebs[65] de la bactérie pour fournir le complexe enzymatique de la nitrogénase en électrons et ATP (figure 35). Par ailleurs, l'apport de phosphate et de sulfate au nodule est aussi un facteur limitant pour le bon développement et fonctionnement du nodule. Là encore, l'importance des transporteurs associés à ces éléments a par exemple été montrée chez le soja (Chen et al., 2019) ou le lotier (Krusell et al., 2005). Au sein du nodule, l'assimilation du NH_3 produit par les bactéroïdes se fait *via* le système GS-GOGAT de la plante à travers la production d'acides aminées, principalement la glutamine, l'alanine et l'acide glutamique (Cordoba et al., 2003). Ces acides aminés peuvent ensuite être exportés du nodule par des transporteurs tels que LjAMT1;1 ou MtNPF7.6 situés au niveau de la membrane plasmique (Rogato et al., 2008; Wang et al., 2020).

Ainsi les principales caractéristiques associées au métabolisme d'un nodule fonctionnel peuvent se résumer en trois points : i) un apport aux bactéroïdes de sources carbonées issues de la photosynthèse, ii) le maintien d'un environnement pauvre en oxygène pour éviter l'oxydation de la nitrogénase et iii) une assimilation par la plante de l'ammonium fourni par les bactéroïdes fixateurs d'azote.

Importance de l'environnement dans l'interaction symbiotique

La symbiose rhizobium-légumineuse bien qu'avantageuse pour la fixation de l'azote atmosphérique, a un coût pour les deux partenaires. Pour faire fonctionner la nitrogénase, la bactérie va utiliser pas moins de 16 molécules d'ATP pour chaque molécule de N_2 que la plante va fournir en remobilisant certaines de ses ressources carbonées. Pour limiter ce coût énergétique des paramètres environnementaux rentrent en jeu pour moduler l'interaction symbiotique.

Une des premières régulations existantes est le contrôle de la nodulation par la plante pour éviter la formation excessive de nodules[1] et cette voie serait principalement activée par des facteurs environnementaux tels que la biodisponibilité en nitrates et phosphates dans les sols (figure 36) (Ferguson et al., 2019). Lorsque le sol n'est pas carencé en azote, la plante va préférentiellement prélever le nitrate et l'ammonium directement par ses racines plutôt que par la fixation biologique à travers la nodulation. En effet la symbiose est contrôlée de manière multiple par la teneur en nitrate des sols (figure 36), notamment par une autorégulation de la nodulation et par une signalisation longue distance à travers les peptides CLE[2] (Oka-Kira & Kawaguchi, 2006) mais aussi par le facteur de transcription NLP[3] (Nishida et al., 2018).

Des stress abiotiques peuvent aussi influencer la mise en place de l'interaction symbiotique comme par exemple la salinité des sols. En effet un stress salin inhibe la nodulation et l'infection par les rhizobiums chez le soja par exemple à travers la modulation de l'expression de NIN (Singh & Valdés-López, 2023). Enfin d'autres mécanismes tels qu'une forte production d'éthylène par la plante, une acidité des sols ou de nombreux autres stress environnementaux impactent la nodulation et sont résumés dans la revue de Ferguson et ses collaborateurs (Ferguson et al., 2019).

[1] En anglais, ce mécanisme est appelé "autoregulation of nodulation" (AON).
[2] CLAVATA/ESR-*related* (CLE) genes, impliqués dans la maintenance des cellules méristématiques.
[3] Découvert à partir du gène NRSYM1 (Nitrate unresponsive symbiosis 1) codant pour une NIN-like protein.

[65] Aussi appelé cycle TCA pour tricarboxilic acid.

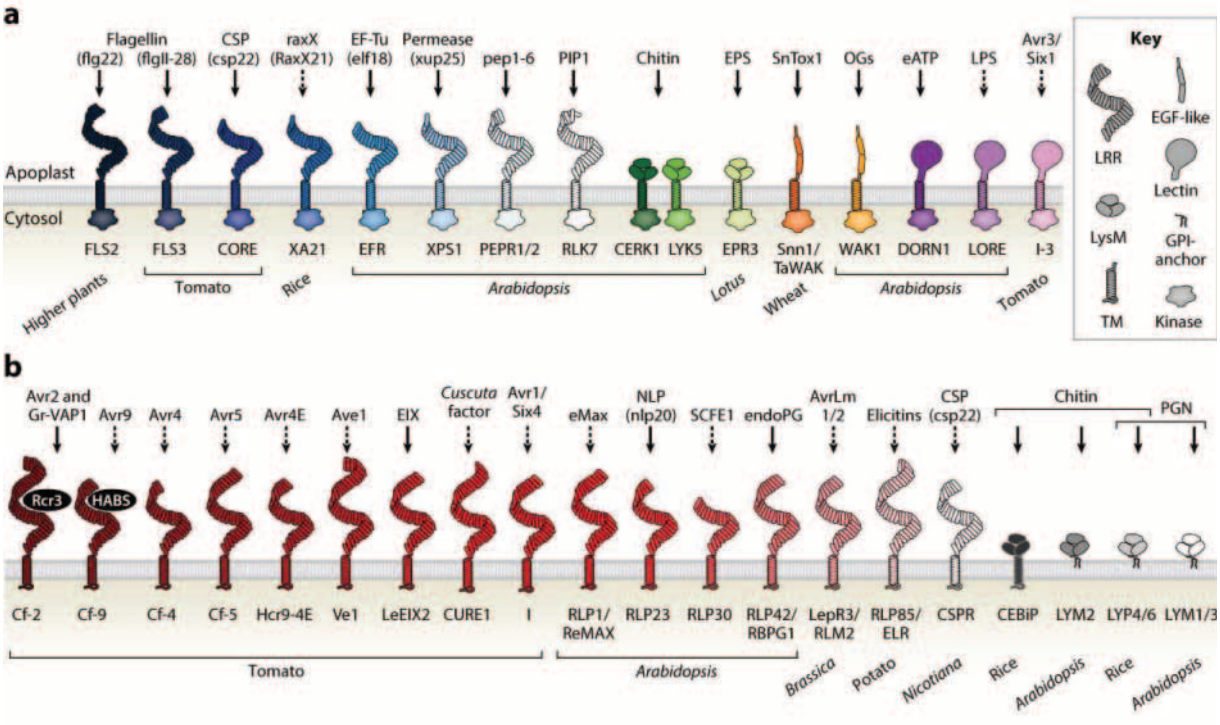


Figure 37. Exemples de récepteurs PRR avec leur ligand associé (Boutrot & Zipfel, 2017).
(**a**) PRR de type RLK (Receptor-Like Kinase). (**b**) PRR de type RLP (Receptor-Like Protein). L'organisme où le récepteur a été découvert est indiqué sous chacun des PRR. Les domaines des PRR sont indiqués dans la légende en haut à droite. Les flèches pleines indiquent une liaison directe démontrée tandis que les flèches en pointillées indiquent l'absence de preuve d'une liaison directe. EGF, facteur de croissance épidermique ; EIX, xylanase induisant l'éthylène ; EPS, exo-polysaccharides de surface ; GPI, glycophosphatidylinositol ; LPS, lipopolysaccharides ; LRR, leucine-rich repeat ; OGs, oligogalacturonides ; PGN, peptidoglycane ; TM, transmembranaire.

4. L'immunité de la plante : un équilibre fragile entre infection et résistance

La diversité microbienne des sols est grande[66] (Delgado-Baquerizo et al., 2018; Torsvik & Øvreås, 2002) et il est donc important pour les rhizobiums de pouvoir être reconnus spécifiquement par la plante et de prendre le dessus par rapport à d'autres microorganismes de la rhizosphère. Parallèlement, la plante doit pouvoir se défendre face à d'éventuels pathogènes tout en permettant l'infection des rhizobiums pour la mise en place de l'organe symbiotique. Ce challenge entre pathogénicité et mutualisme va être discuté dans les parties qui vont suivre.

4.1. Généralités sur l'immunité de la plante : contexte d'une plante résistante

Dans son environnement, la plante doit pouvoir se défendre face à des organismes pathogènes incluant des bactéries, des champignons, des virus, des nématodes mais aussi des insectes. A l'inverse des animaux, elles n'ont pas développé un système immunitaire adaptatif. Cependant, chaque cellule végétale est capable de déclencher une réponse immunitaire dite innée conférant à la plante la capacité de répondre immédiatement et de manière non-spécifique en cas d'agression par un agent pathogène. Connaître les mécanismes de défense existants chez les plantes est crucial pour mieux comprendre comment l'interaction symbiotique rhizobium-légumineuse est possible.

4.1.1. Immunité à la surface des cellules : première ligne de défense

Le premier niveau de perception des microorganismes par la plante est basé sur la reconnaissance de molécules microbiennes conservées au cours de l'évolution appelés MAMPs ou PAMPs[67]. Cette reconnaissance du pathogène ou encore l'induction d'une blessure, entraîne aussi la production par la plante d'autres molécules pouvant être reconnues comme signaux de danger et appelés DAMPs[68] (Schwessinger & Ronald, 2012). La perception de l'ensemble de ces molécules nécessite l'existence de récepteurs localisés sur la membrane plasmique nommés PRR[69] (Boutrot & Zipfel, 2017; Zipfel, 2009). Les récepteurs PRR sont des RLK[70] ou des RLP[71] qui diffèrent par la présence ou l'absence d'un domaine kinase pour la transduction du signal (figure 37). Ils contiennent tous des domaines extra-

[66] On estime qu'un gramme de sol renferme environ un milliard de bactéries et entre 2 000 et 10 000 espèces bactériennes (Roesch et al., 2007).
[67] Microbe-Associated Molecular Patterns ou Pathogen-Associated Molecular Patterns (peptidoglycane, flagelline, FNs etc.). Dénomination proposée en 2004, PAMPs étant l'ancien nom qui a été renommé en MAMPs en raison du caractère non pathogène d'un dérivé de peptidoglycane impliqué dans la signalisation symbiotique entre le calamar *Euprymma scolopes* et la bactérie *Vibrio fischeri* (Koropatnick et al., 2004).
[68] Damage-Associated Molecular Patterns : fragments de paroi cellulaires ou extracellulaires, peptides, nucléotides, acides aminés etc. (Hou et al., 2019).
[69] Pattern Recognition Receptors.
[70] Receptor-Like Kinases.
[71] Receptor-Like Proteins.

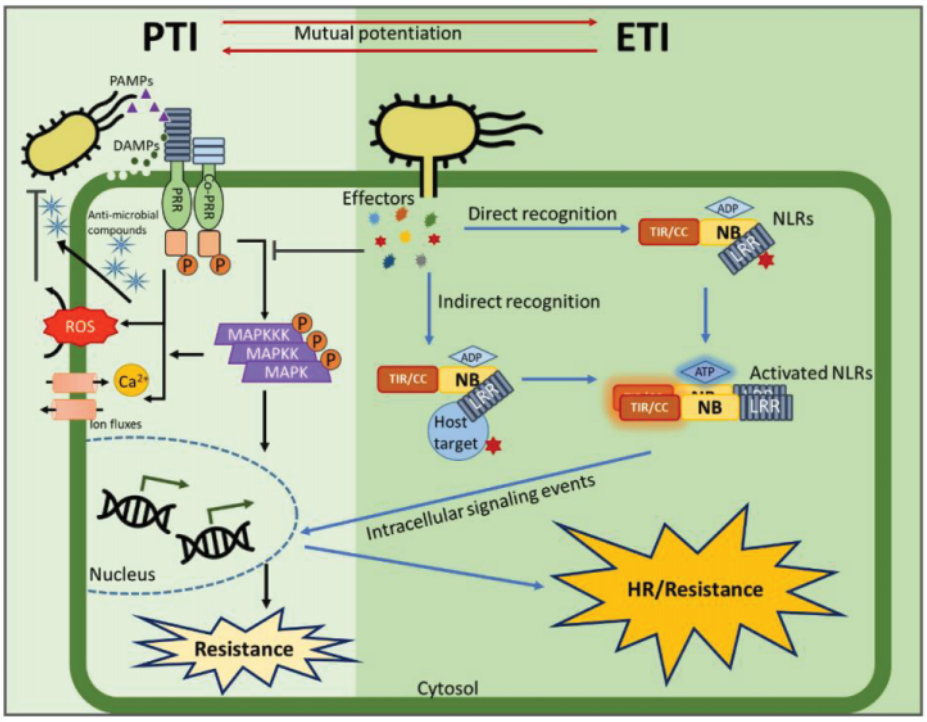


Figure 38. Représentation schématique de l'immunité chez les plantes (Nguyen et al., 2021).
La PTI, première ligne de défense, est déclenchée par la reconnaissance de motifs moléculaires (PAMPs, MAMPs et DAMPs) par l'intermédiaire de récepteurs PRR. Plusieurs évènements de signalisation sont associés à la PTI : activation de la voie des MAPK (Mitogen-Activated Protein Kinases), influx de Ca^{2+} dans le cytoplasme et production de ROS. Des composés antimicrobiens sont également produits contre le pathogène et les gènes de défense sont activés dans le noyau. Au cours de l'évolution, les pathogènes ont réussi à surmonter cette première ligne défense à travers la sécrétion d'effecteurs directement dans la cellule hôte. En réponse à cette invasion, la plante a su développer une seconde ligne de défense, l'ETI, à travers des protéines de résistance (NBS-LRR) capables de reconnaître spécifiquement les effecteurs (de manière directe ou indirecte). Cette reconnaissance active d'autres évènements de signalisation qui aboutissent à une réaction hypersensible (HR) de la plante ainsi qu'une amplification globale des défenses déjà mises en place durant la PTI.

cellulaires ; principalement des domaines LRR[72] qui sont capables de reconnaître des protéines et des peptides, ou des domaines LysM[73] capables de reconnaître des polysaccharides. La reconnaissance des signaux microbiens par les PRR fonctionne probablement en complexes de différents corécepteurs (Roux et al., 2011; Tang et al., 2017) qui permettent par exemple une amplification du signal ou une stabilisation des récepteurs PRR (Oldroyd, 2013). Par ailleurs, les PRR tout comme les récepteurs des FNs et des EPS de rhizobium, peuvent contenir des domaines LysM extracellulaires, ce qui souligne la mince frontière existante entre la reconnaissance d'un pathogène à celle d'un symbiote bien que les deux voies de transduction du signal semblent être indépendantes (Bozsoki et al., 2017).

Suite à la reconnaissance de ces molécules, la plante active sa première barrière de défense appelée PTI[74] (Nguyen et al., 2021). Celle-ci passe par l'activation d'une cascade de protéines kinases, appelée voie des MAPK[75], qui par phosphorylations successives va activer l'expression des gènes de défense de la plante (figure 38). En parallèle, la production de ROS par les peroxydases permet aussi d'agir de manière directe contre le pathogène bien que leur rôle respectif chez les plantes ait été peu étudié et surtout décrit chez *Arabidopsis thaliana* (Rasmussen et al., 2012; Tang et al., 2017). Parmi les réponses cellulaires induites, on peut noter la production de différentes molécules de défenses dans le but de ralentir et de stopper la progression du pathogène. Cela passe entre autres par la synthèse de composés structuraux telle que la callose, la sécrétion de phytoalexines aux propriétés antimicrobiennes, la synthèse de protéines de résistances[76] ainsi que celle d'inhibiteurs de protéases (Zhou & Zhang, 2020). La production d'hormones de défense, telles que l'acide salicylique, l'éthylène et l'acide jasmonique, intervient dans la modulation de la réponse immunitaire, avec une importance centrale de l'acide salicylique dans l'amplification des mécanismes de défense (Zhou & Zhang, 2020). Enfin, les facteurs de transcription WRKY favorisent l'induction des gènes nécessaires à l'immunité déclenchée par la reconnaissance des PAMPs (Wani et al., 2021). Cette réponse immunitaire dite "basale " est généralement suffisante pour contenir l'infection du pathogène.

4.1.2. Immunité intracellulaire : un second niveau de défense en réponse à l'infection

Cependant, des pathogènes efficaces ont acquis la capacité à surmonter la PTI à travers la sécrétion dans la cellule hôte de molécules capables de supprimer le premier niveau de défense de la plante, appelées effecteurs (figure 38). La sécrétion de telles molécules entraîne une sensibilité chez la plante appelée ETS[77] et peut conduire à une infection chronique si la plante ne peut plus se défendre.

[72] Leucine Rich-Repeat.
[73] Lysine Motif.
[74] Pattern-Triggered Immunity.
[75] Mitogen-Activated Protein Kinases.
[76] Aussi appelées PR protéines pour Pathogen-Related.
[77] Effector-Triggered Susceptibility.

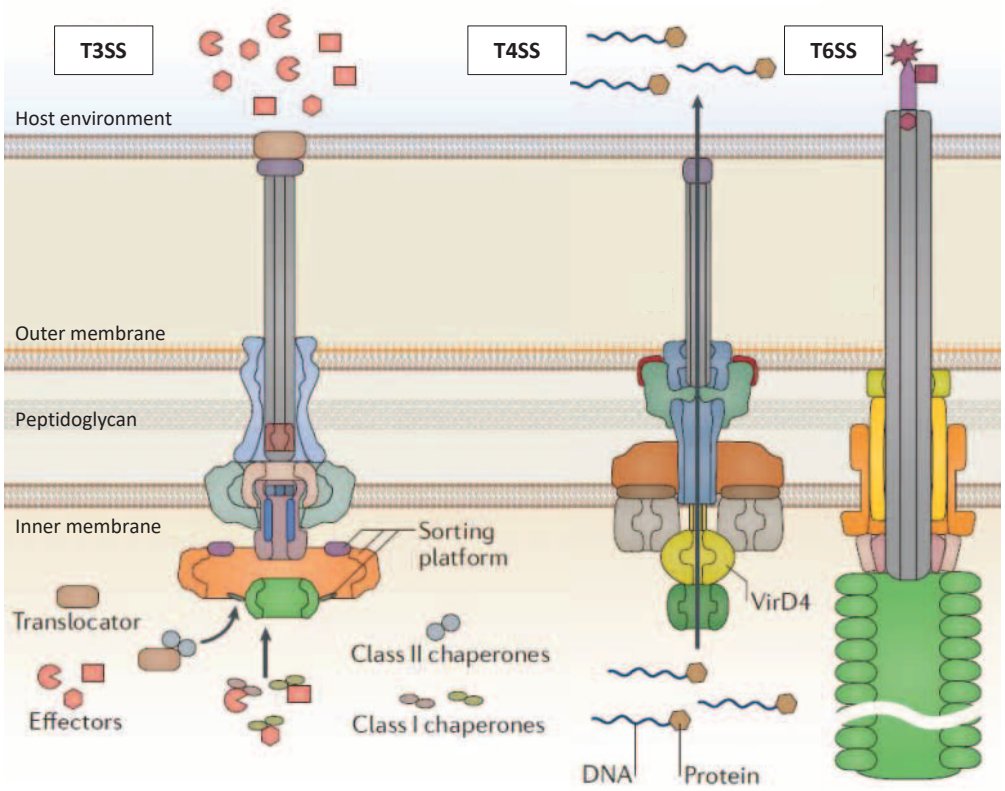


Figure 39. Représentation schématique des trois systèmes de sécrétions utilisés par les bactéries Gram- pour sécréter des effecteurs dans la cellule hôte (adapté de Costa et al., 2015).
Pour contourner la PTI, les bactéries pathogènes ont développées des systèmes de sécrétion de type 3, 4 ou 6 (T3SS, T4SS et T6SS) pour assurer une infection chronique à travers l'envoi de protéines effectrices dans la cellule hôte. Ces systèmes aussi appelés injectisomes sont assimilables à une nano-seringue capable de percer la paroi et la membrane plasmique de la cellule végétale.

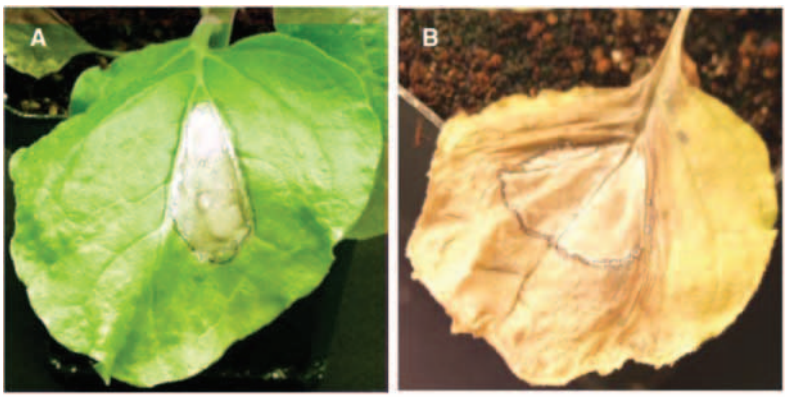


Figure 40. Réaction hypersensible chez *Nicotiana benthamiana* (Balint-Kurti, 2019).
(**A**) La présence du gène N confère une HR en présence du virus TMV (tobacco mosaic virus). La maladie ne se propage pas et le pathogène est stoppé. (**B**) Un gène nécessaire à l'autophagie (*NbBECLIN*) a été rendu silencieux suite à l'infection du virus VIGS ce qui entraîne une propagation de la HR à partir du site initial de son activation.

Les effecteurs sont délivrés dans le cytoplasme de la plante à travers deux différents types de systèmes de sécrétion chez les bactéries Gram-négatives, le T3SS et le T4SS, et il peut être supposé que des effecteurs soient également délivrés par le T6SS (figure 39) (Costa et al., 2015). Ces structures protéiques sont assimilables à une nano-seringue et sont retrouvées autant chez les pathogènes d'animaux que chez les pathogènes de plantes[78]. Les effecteurs délivrés spécifiquement par le T3SS sont alors appelés effecteurs de type 3 (T3Es) ; la même dénomination peut être utilisée pour les T4Es et T6Es. Il est intéressant de noter que les champignons sécrètent eux aussi des effecteurs, à travers des structures intracellulaires (hyphes, haustorium et arbuscules), qui jouent également un rôle dans la pathogénicité du champignon (Lo Presti et al., 2015). Face à la virulence du pathogène capable de sécréter des effecteurs, un second niveau de perception est apparu chez la plante au cours de l'évolution, appelé ETI[79], qui cette fois fait intervenir des récepteurs intracellulaires. Cette réponse repose sur la reconnaissance spécifique d'un effecteur par le produit d'un gène de résistance qui possède dans la majorité des cas des sites de liaisons nucléotidiques spécifiques et des domaines répétés riches en leucine, appelés récepteurs NBS-LRR[80] (Cui et al., 2015; Jones & Dangl, 2006). L'ensemble de ces gènes de résistance code pour des protéines de résistance aussi appelées R protéines. Des réponses de défenses similaires à la PTI sont alors mises en place mais ces dernières sont plus robustes et prolongées dans le temps et conduisent à la mort cellulaire programmée des tissus infectés pour stopper la prolifération du pathogène : on parle de réaction hypersensible[81] (Katagiri & Tsuda, 2010). Cette réponse HR localisée au niveau du site de reconnaissance d'un effecteur donné (figure 40), est associée à une production massive de ROS et de composés antimicrobiens (Balint-Kurti, 2019). L'effecteur ainsi reconnu est responsable de l'incompatibilité d'hôte entre le pathogène et la plante, conduisant à la résistance de cette dernière : on parle alors de facteur d'avirulence[82] pour décrire l'effecteur reconnu. Enfin, ces gènes de résistance sont des éléments fondamentaux de la reconnaissance des effecteurs intracellulaires et font l'objet d'une sélection variée pour générer de nouvelles surfaces protéiques adaptées à la reconnaissance de nouveaux effecteurs, mais ils sont aussi des gardiens de l'homéostasie cellulaire qui surveillent les signaux de danger (Cui et al., 2015).

[78] Le rôle du T3SS chez les rhizobiums a fait l'objet d'une revue disponible dans la partie 6 de l'introduction.
[79] Effector-Triggered Immunity.
[80] Nucleotide Binding Site-Leucine-Rich Repeats ou NLR. Il existe 3 grandes familles selon leur partie N-terminale : les TIR-NBS-LRR (TNL) possédant un domaine Toll Interleukin-1 Receptor, les CC-NBS-LRR (CNL) avec un domaine Coiled-coil et les RPW8-NB-ARC-LRR (RNL) portant un domaine RPW8-like (Zhou & Zhang, 2020).
[81] Aussi appelée réponse HR pour "hypersensitive response".
[82] Se dit d'un effecteur lorsqu'il perd son pouvoir pathogène et que la maladie disparaît.

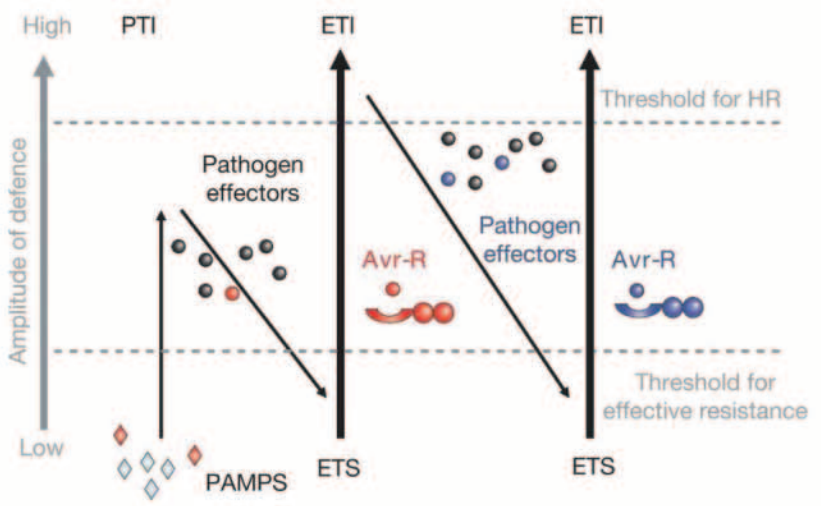


Figure 41. Modèle en zig-zag illustrant la coévolution hôte-pathogène (Jones & Dangl, 2006).
L'adaptation du pathogène pour supprimer la PTI a conduit à la sécrétion d'effecteurs. Ces derniers confèrent la maladie à la plante jusqu'à ce qu'ils soient reconnus par des récepteurs NLR issus de l'adaptation de la plante au pathogène. Ce processus peut ainsi se répéter et être assimiler à une "course à l'armement".

4.1.3. Coévolution des gènes de résistance et des effecteurs : le modèle en zig-zag

La virulence d'un effecteur peut être restaurée soit par sa mutation qui empêche la reconnaissance avec la protéine de résistance associée, soit par l'émergence de nouveaux effecteurs virulents. A ce moment-là, il y a perte de l'ETI et l'interaction devient compatible (la maladie s'installe à nouveau). La plante peut alors à son tour riposter en produisant une nouvelle R protéine qui reconnaîtra à nouveau l'effecteur responsable de la pathogénicité. Cette suppression et restauration consécutive de l'ETI souligne une pression de sélection antagoniste entre l'acquisition de nouveaux gènes de résistances et celle de nouveaux effecteurs virulents. La résistance d'une plante semble donc avoir une efficacité limitée dans le temps jusqu'à l'apparition d'un nouvel effecteur et inversement. Ce modèle peut être comparable à une "course à l'armement", où le caractère pathogène du microorganisme et résistant de l'hôte est fluctuant dans le temps (figure 41). Cette notion a été décrite pour la première fois sous le nom de modèle en zig-zag (Jones & Dangl, 2006) et souligne finalement la fragilité d'une résistance monogénique (ETI) bien que plus performante qu'une résistance basale (PTI).

4.2. Spécificité d'interaction

Le fait de posséder un set de gènes essentiels à la symbiose ne garantit pas une interaction symbiotique compatible avec toutes les légumineuses existantes. En effet chaque souche de rhizobium possède son spectre d'hôte de légumineuse(s) spécifique(s) et inversement. Ce spectre d'hôte peut être plus ou moins large et dépend à la fois de la plante et de la bactérie (Andrews & Andrews, 2017). Cette sélection de l'hôte et du symbiote met en jeu l'immunité de la plante et est principalement médiée par la sécrétion de flavonoïdes (cf. 3.3.1) mais aussi dans la plupart des cas par la production de FNs synthétisés par les rhizobiums (cf. 3.3.2) ainsi que par la sécrétion d'effecteurs.

4.2.1. Spectre d'hôte : les déterminants bactériens

La plupart des rhizobiums ont un spectre d'hôte étroit. Par exemple, *Rhizobium leguminosarum* bv. *trifolii* est capable d'infecter uniquement les espèces de trèfle (*Trifollium*) (Poole et al., 2018). Par opposition, il existe des rhizobiums qui ont un spectre large comme *Sinorhizobium fredii* NGR234 capable d'infecter 112 espèces de légumineuses en plus de la non légumineuse *Parasponia andersonii*.

Une des principales étapes pour discriminer les différentes espèces de bactéries et sélectionner le bon symbiote, s'effectue lors du développement et de la propagation du cordon d'infection. A ce stade, les FNs jouent un rôle clé dans la spécificité d'hôte. Leur variabilité permet aux rhizobiums d'être reconnus spécifiquement et d'induire des activités biologiques à la fois sur des légumineuses

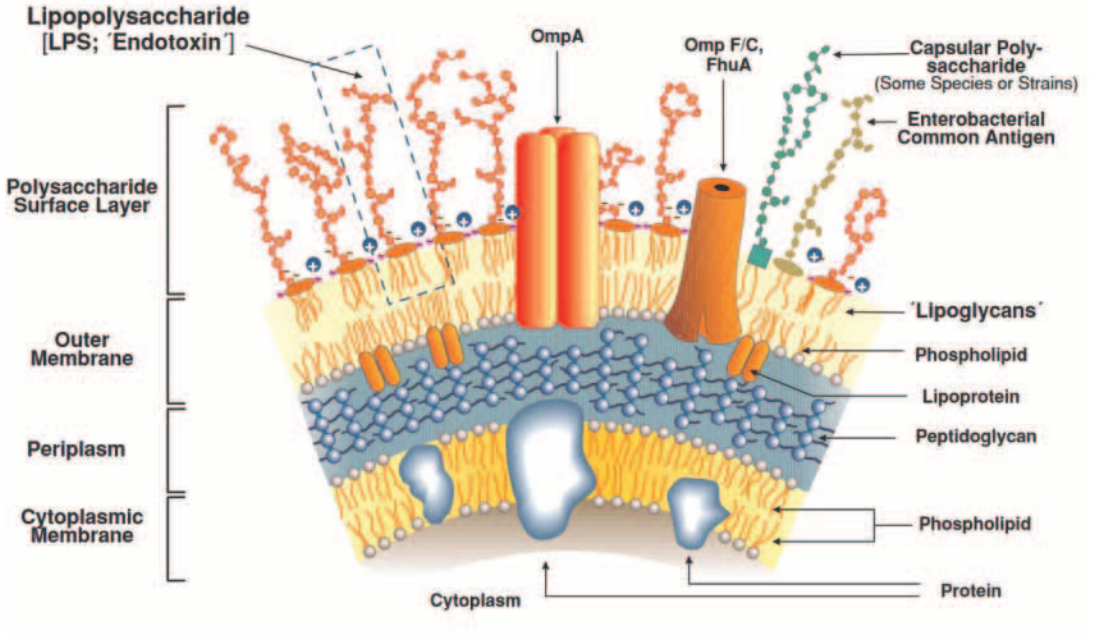


Figure 42. Architecture de la paroi cellulaire des bactéries Gram-négatives (Alexander & Rietschel, 2001).
Les rhizobiums sont des bactéries Gram-négatives caractérisées par la présence de deux bicouches lipidiques. La membrane externe de la bactérie est composée de régions lipidiques composées de LPS qui représentent environ 75% de la surface membranaire. Ces LPS peuvent rentrer en contact avec des protéines membranaires telles que les porines OmpA, OmpF/C ou le récepteur FhuA. Certaines souches possèdent aussi des KPS (Capsular Polysaccharides). Les EPS (exo-polysaccharides) ne sont pas représentés sur ce schéma car ils sont sécrétés dans le milieu extracellulaire. Ils permettent entre-autres de masquer la présence de la bactérie facilement reconnaissable par la présence des LPS et KPS.

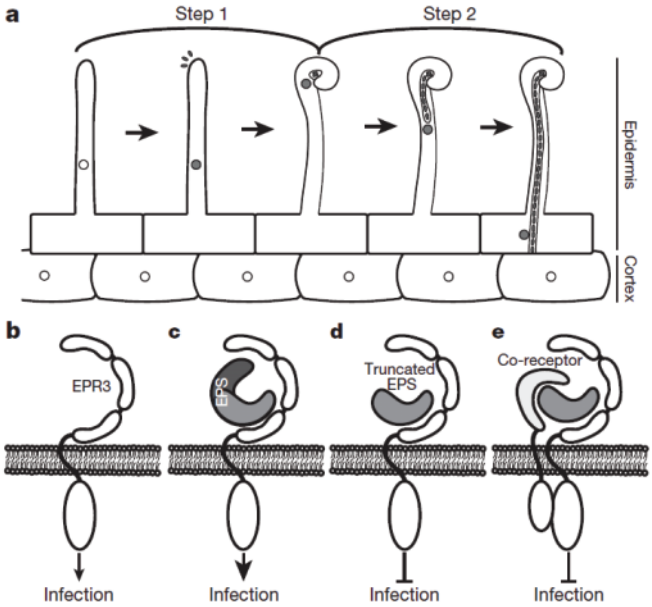


Figure 43. Modèle de reconnaissance en deux étapes des rhizobiums compatibles par l'intermédiaire du récepteur EPR3 (Kawaharada et al., 2015).
(**a**) La reconnaissance des facteurs Nod entraîne l'expression d'EPR3 dans les cellules épidermiques (étape 1) puis la voie de signalisation se poursuit et EPR3 reconnaît les EPS (étape 2). (**b-c**) L'infection est possible lors d'un niveau basal d'EPR3 ou lorsqu'il perçoit un EPS compatible. (**d-e**) Un EPS tronqué ou incompatible active une voie bloquant l'infection et ce blocage peut être médié par un co-récepteur.

tropicales et sur des légumineuses tempérées (Price et al., 1992). Il a ainsi pu être montré que des changements du domaine extracellulaire des récepteurs aux FNs induisent une augmentation du spectre d'hôte chez deux espèces de *Lotus* (Radutoiu et al., 2007), soulignant l'importance de ces molécules à être reconnues par la plante pour initier la symbiose. D'autres facteurs bactériens sont aussi déterminants pour le spectre d'hôte : ce sont les polysaccharides de surface[83] (Alexander & Rietschel, 2001) (figure 42). Ils sont principalement responsables de la protection de la bactérie contre tout type de stimuli pouvant provenir directement du sol, mais aussi de la rhizosphère et de la plante elle-même. Il a pu être montré que ces composés de surfaces, la plupart du temps reconnus négativement par la plante dans un contexte pathogène, ont évolué chez les rhizobiums de manière à adapter la réponse immunitaire de la plante au cours de la nodulation (Gourion et al., 2015). En plus de cette modulation de l'immunité de l'hôte, les polysaccharides de surface sont aussi indispensables au cours de l'interaction pour leur rôle dans l'adhésion de la bactérie aux racines (Downie, 2010).

Enfin, en raison de leur grande diversité de structures au sein des rhizobiums, il est plausible que la reconnaissance des polysaccharides puisse participer à la spécificité d'hôte tout comme les FNs (Walker et al., 2020). L'identification d'un récepteur LysM (EPR3[84]) chez *L. japonicus* capable de reconnaître des exopolysaccharides suite à la perception des FNs vient appuyer cette hypothèse (figure 43) (Kawaharada et al., 2015). Lorsqu'on s'intéresse à la capacité de la souche NGR234 à développer une symbiose efficiente avec autant de légumineuses, on relève que le spectre d'hôte peut aussi être médié par les systèmes de sécrétions bactériens. Chez NGR234, deux T3SS et un T4SS ont été montrés comme influant le spectre d'hôte (Schmeisser et al., 2009). Ces systèmes sont connus dans la symbiose rhizobienne pour moduler la réponse immunitaire de la plante à travers la sécrétion de protéines effectrices[85], qui peuvent promouvoir la symbiose ou au contraire l'abolir[86]. Des études ont également montré que la reconnaissance des effecteurs (T3Es ou T4Es) dans la cellule hôte est un moyen de sélection des symbiotes, comme cela a pu être montré chez le soja ou le niébé (Yang et al., 2010; Hayashi et al., 2011; Nguyen et al., 2017). Ces différentes molécules (FNs, polysaccharides et effecteurs) peuvent donc être reconnues négativement et provoquer la mise en place de mécanismes de défense par la plante dans le cas d'une interaction incompatible. C'est donc grâce à un équilibre entre l'immunité de la plante, qui régule la sélection du microbiote pouvant varier selon l'environnement, avec la modulation de cette réponse immunitaire par les rhizobiums que l'infection peut se faire (Sprent et al., 2017; Walker et al., 2020).

[83] On peut citer par exemple les LPS (lipopolysaccharides), les EPS (exo-polysaccharides) ou encore les KPS (K-antigen polysaccharide ou polysaccharides capsulaires).
[84] Exopolysaccharide receptor 3.
[85] A l'inverse du T3SS, le T4SS est aussi capable de transporter et de sécréter de l'ADN.
[86] Les rôles spécifiques des T3Es sont détaillés dans la revue disponible dans la partie 6 de cette introduction.

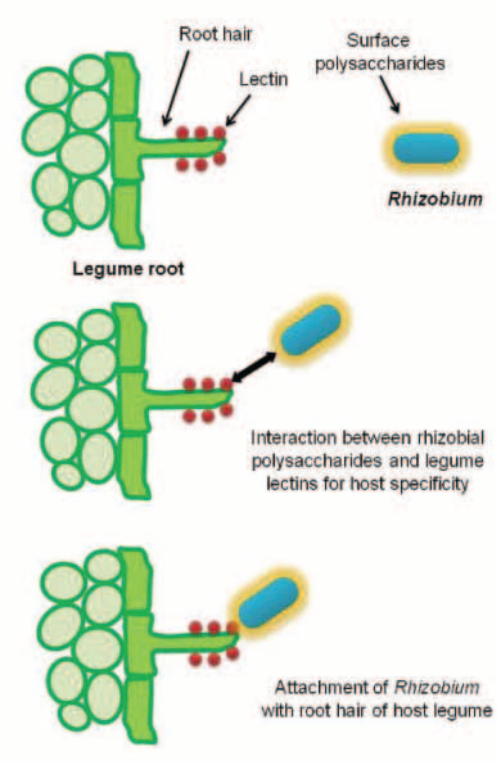


Figure 44. Modèle de reconnaissance des exo-polysaccharides des rhizobiums par les lectines (Ahemad & Khan, 2013).
Les lectines produites par les légumineuses peuvent reconnaitre de manière spécifique certains EPS de rhizobium, influençant ainsi le spectre d'hôtes. Les rhizobiums ayant adhérés au poil racinaire pourront ensuite poursuivre le processus d'infection.

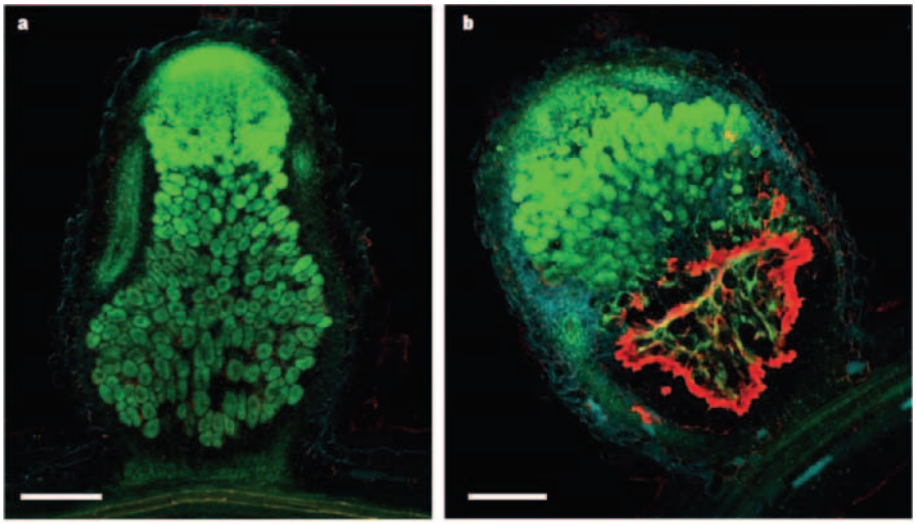


Figure 45. La nodulation nécessite une atténuation des défenses de la plante (Zipfel & Oldroyd, 2017).
Nodule racinaire de *M. truncatula* (a) et d'un mutant 7Y de *M. truncatula* (b) incapable de restreindre ses mécanismes de défense dans les cellules infectées par les rhizobiums. Les acides nucléiques de la plante et des bactéries sont colorés en vert par le SYTO13. Les composés phénoliques caractéristiques de l'activation de l'immunité de la plante auto-fluorescent en rouge. Echelle : 25 mm.

4.2.2. Spectre d'hôte : les déterminants plantes

Parallèlement, le contrôle par la plante des souches capables de pénétrer leurs tissus est crucial pour éviter les bactéries opportunistes qui ne vont par leur apporter de bénéfice (Walker et al., 2020). Tout comme certains rhizobiums sont capables d'infecter un grand nombre d'espèces de légumineuses, il existe aussi des légumineuses capables d'être infectées par une large diversité de rhizobiums[87]. C'est le cas par exemple de l'espèce de niébé *Vigna unguiculata* qui est nodulée par une grande diversité de *Bradyrhizobium* et quelques bactéries du genre *Rhizobium* (Ndungu et al., 2018) ; ou encore *Lotus burttii* qui est nodulée par cinq genres différents de rhizobium (Zarrabian et al., 2022). Parmi les déterminants plantes rentrant en jeu pour la spécificité d'hôte, les flavonoïdes sont les premières molécules à médier le spectre d'hôte. Leur grande diversité participe à ce contrôle et influence la gamme d'hôte en agissant soit comme un signal d'infection qui stimule l'expression des gènes *nod* bactériens à travers leur reconnaissance par le facteur de transcription NodD (interaction compatible), soit en agissant comme une phytoalexine et en déclenchant une activité antimicrobienne (interaction incompatible) (Liu & Murray, 2016).

La production de lectines[88] par les plantes est un autre déterminant du spectre d'hôte qui favorise l'adhésion des rhizobiums au tissu végétal en se liant à des polysaccharides de surface spécifiques (Laus et al., 2006; Smit et al., 1992). La plante hôte favorise ainsi l'attachement de certaines bactéries par rapport à d'autres pour mieux contrôler le choix de son partenaire symbiotique (Ahemad & Khan, 2013) (figure 44). Par la suite, la plante réprime ses mécanismes de défense pour tolérer intracellulairement la bactérie pouvant éventuellement conduire à la fixation d'azote (Gourion et al., 2015). Cependant si les réactions de défense à l'intérieur du nodule ne sont pas contrôlées, ce dernier devient nécrotique. Dans ce sens, un mutant plante incapable de restreindre ses mécanismes de défense empêche la bonne mise en place de la symbiose et voit son nodule avorté (figure 45). Dans le cas particulier des légumineuses appartenant à la famille des Papilionoideae, il peut être observé une régulation de la gamme d'hôtes par des NCR. Cette grande famille de peptides antimicrobiens montre une expression associée au développement du nodule et certains NCR peuvent restreindre la symbiose avec des souches de rhizobiums spécifiques tandis que d'autres sont essentiels (Walker et al., 2020). Il a été suggéré que la tolérance des rhizobiums aux NCR dépend de l'activité de protéines de type BacA[89] qui serait médiée par leur capacité à interagir avec ces peptides antimicrobiens. De plus, il a pu être montré que certains NCR peuvent aussi influencer la viabilité des bactéroïdes,

[87] On parle alors de "promiscuous legumes" pouvant se traduire par "légumineuses non-sélectives".

[88] Les lectines sont des protéines synthétisées par les plantes capables de se lier spécifiquement à certains glucides. Elles interviennent dans plusieurs processus biologiques ainsi que dans l'attachement et la liaison de divers micro-organismes.

[89] Protéines de transport membranaire nécessaires à la survie des rhizobiums dans le symbiosome (Roy et al., 2020).

comme NCR343 chez *M. truncatula* (Gao et al., 2023), renforçant le pouvoir de la plante face au rhizobium.

Enfin bien que la sélection du symbiote ait généralement lieu durant les étapes précoces de l'interaction symbiotique, il existe des mécanismes pour discriminer les rhizobiums qui ne seraient pas bénéfiques pour la plante une fois le nodule formé. C'est ce qu'on appelle des sanctions de l'hôte qui discrimine et régule le fonctionnement du nodule une fois la symbiose établie (Walker et al., 2020). Ainsi, la plante peut contrôler la viabilité des rhizobiums dans les nodules et "sanctionner" les souches peu performantes ou tricheuses *via* une régulation de l'allocation des ressources carbonées vers le nodule (Westhoek et al., 2017, 2021). Ce processus souligne le contrôle important de la plante sur ses partenaires symbiotiques. Il n'a cependant pas été observé chez un grand nombre de légumineuses et a souvent été démontré de manière artificielle et non en condition naturelle (Walker et al., 2020). C'est finalement un contrôle réciproque des deux partenaires symbiotiques qui a lieu puisqu'autant les rhizobiums que les légumineuses peuvent restreindre ou promouvoir leur gamme d'hôtes pour la fixation d'azote. Néanmoins aucun lien entre la phylogénie des légumineuses et leur rhizobium(s) associé(s) n'a pu être établie à ce jour, soulignant le caractère complexe de cette spécificité d'hôte (Doyle, 1998; Martínez-Romero, 2009; Rahimlou et al., 2021).

Immunité de la plante dans un contexte symbiotique

Durant la symbiose rhizobium-légumineuse, les défenses doivent être supprimées ou évitées par les rhizobiums pour réussir à infecter les cellules jusqu'à atteindre le nodule. En réponse à la bactérie, il a pu être montré qu'il y a une induction transitoire des gènes de défense aux premiers stades de l'interaction symbiotique ; mais que par la suite le niveau d'expression des gènes de défense revient à un niveau basal ou est même réprimé pour favoriser l'infection (Libault et al., 2010; Samac et al., 2006).

Diverses stratégies ont été mises en place par la plante ou son symbiote pour favoriser la compatibilité de l'interaction. Parmi elles, la répression de l'éthylène est un moyen utilisé par les bactéries pour limiter les mécanismes de défense de la plante en faveur de la nodulation. En effet, lors de la PTI cette hormone de défense est produite majoritairement dans les racines et impacte négativement le nombre de nodules, suggérant que l'éthylène réprime la nodulation. Pour coloniser la plante, les rhizobiums tout comme les microorganismes pathogènes utilisent des protéines effectrices qui sont transférées dans les cellules végétales pour modifier diverses fonctions cellulaires et faciliter la colonisation et l'infection (Deakin & Broughton, 2009). Cependant cette stratégie est à double-tranchant puisque comme pour les effecteurs pathogènes, ceux associés à la symbiose peuvent aussi être reconnus et rendre l'interaction incompatible (Staehelin & Krishnan, 2015; Wang et al., 2017). Par ailleurs, la production de peptides antimicrobiens (NCR[1]) chez certaines légumineuses pour garder le contrôle sur la bactérie souligne le recrutement de l'immunité durant l'interaction symbiotique. Enfin, la production de signaux calciques tout comme celle de ROS durant la symbiose, démontre une activation du système immunitaire qui recrute les mêmes acteurs mais à des niveaux d'intensité différents. L'ensemble de ces observations indique que des acteurs communs au système immunitaire sont recrutés durant la symbiose fixatrice d'azote rhizobium-légumineuse et que leur contrôle est nécessaire pour aboutir à une symbiose efficiente.

[1] Nodule-specific Cysteine-Rich peptides.

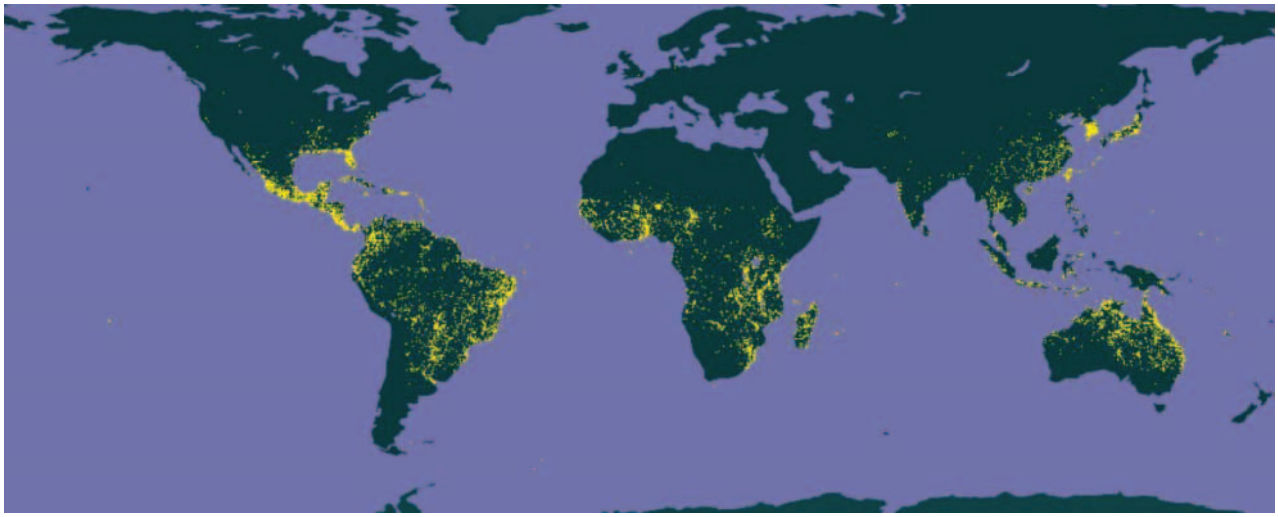


Figure 46. Répartition des différentes localisations géographiques où des *Aeschynomene* spp. ont été identifiées (GBIF). Chaque point jaune représente un lieu où des *Aeschynomene* ont été identifiées.

5. Existence de symbioses nodulaires atypiques

Dans les parties précédentes, nous avons décrit principalement les processus morphologiques et moléculaires conduisant à la symbiose chez les légumineuses modèles *L. japonicus* et *M. truncatula*. Chez ces deux légumineuses, les FNs jouent un rôle central à la fois dans le contrôle de l'infection de la plante par les rhizobiums mais aussi dans le développement de nodules racinaires. Cependant, les FNs ne constituent pas un prérequis pour toutes les symbioses rhizobium-légumineuse puisque des exceptions existent. C'est le cas de certaines espèces du genre *Aeschynomene* qui peuvent être nodulées de manière Nod-indépendante par différents types de *Bradyrhizobium* photosynthétiques ou non, possédant ou pas un T3SS et qui possèdent un certain nombre de particularités atypiques qui vont être décrites dans les paragraphes qui vont suivre.

5.1. Le couple symbiotique *Bradyrhizobium*/*Aeschynomene* spp.

5.1.1. Les *Aeschynomene* : généralités

Les plantes du genre *Aeschynomene*[90] font partie du clade des Dalbergioïdes qui a été décrit en 2001 par Lavin et ses collaborateurs et qui se compose majoritairement d'espèces tropicales et subtropicales. C'est un groupe monophylétique qui comprend 44 genres et environ 1100 espèces d'arbres, arbustes et plantes vivaces, réparties dans trois sous-clades que sont les Adesmia, Dalbergia et Pterocarpus (Lavin et al., 2001). Le groupe des Dalbergioïdes se caractérise entre autres par la capacité à former des nodules d'une morphologie unique de type "aeschynoménoïde" (mise à part quatre exceptions que sont *Andira*, *Hymenolobium*, *Vatairea* et *Vataireopsis* - Lavin et al., 2001). Un nodule est dit "aeschynoménoïde" lorsque l'infection a lieu à l'émergence des racines latérales ou adventives (dans le cas des nodules de tige), plutôt que par l'invagination de la membrane cytoplasmique d'un poil racinaire. Ce sont des nodules sphériques à croissance déterminée mais leur structure interne est bien différente des nodules infectés par cordon d'infection (Sprent et al., 2013). La zone d'infection du nodule est en effet homogène et l'ensemble des cellules du tissu central est infecté intracellulairement. Ce type de morphologie est possible par la division d'une cellule ou quelques cellules fondatrices initialement infectées qui vont se diviser successivement permettant d'obtenir ce type d'infection homogène (Bonaldi et al., 2011; Lavin et al., 2001).

La diversité connue des *Aeschynomene* s'élève entre 175 à 180 espèces mais il est parfois estimé que ce nombre pourrait atteindre les 250 spécimens (Lewis et al., 2005). Une grande majorité des représentants de ce genre se situe dans les pays néo et subtropicaux (figure 46), tels que le Mexique et

[90] Dérivé du latin "aeschynomeno", ancien nom désignant une plante sensible en référence à la sensitivité des feuilles d'*Aeschynomene* au toucher et à la température.

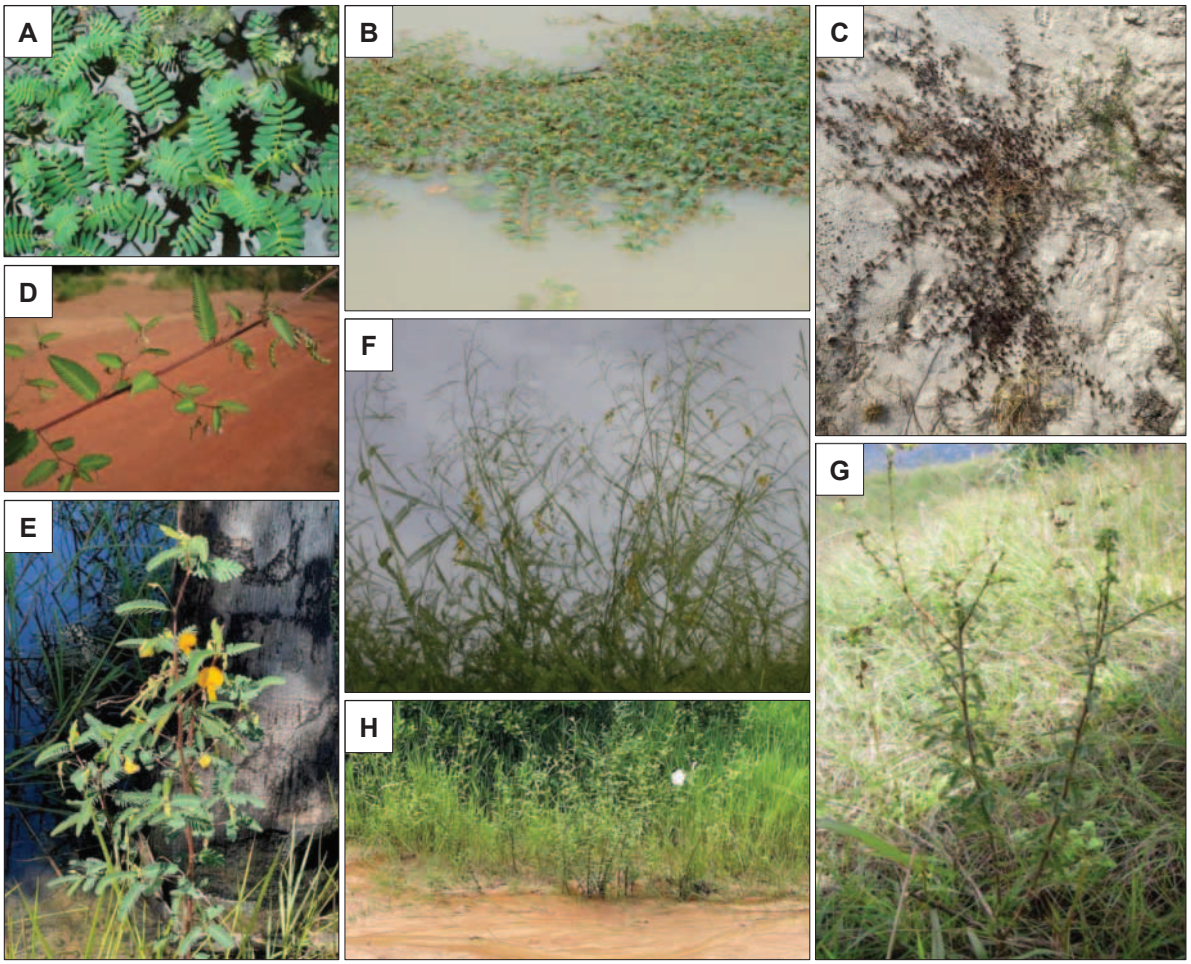


Figure 47. Diversité des environnements où se développent *Aeschynomene* spp.
(**A**) *Aeschynomene fluitans* (Wikipedia). (**B**) *Aeschynomene crassicaulis* (Arne Erpenbach – Benin). (**C**) *Aeschynomene patula* (Jean-François Arrighi – Madagascar). (**D**) *Aeschynomene americana* (Marco Schmidt – Burkina Faso). (**E**) *Aeschynomene cristata* (Jean-François Arrighi – Madagascar). (**F**) *Aeschynomene nilotica* (Philippe Birnbaum – Mali). (**G**) *Aeschynomene abyssinica* (David Schellenberger Costa – Tanzanie). (**H**) *Aeschynomene sensitiva* (gauche) et *Aeschynomene indica* (droite) (Jean-François Arrighi – Madagascar).

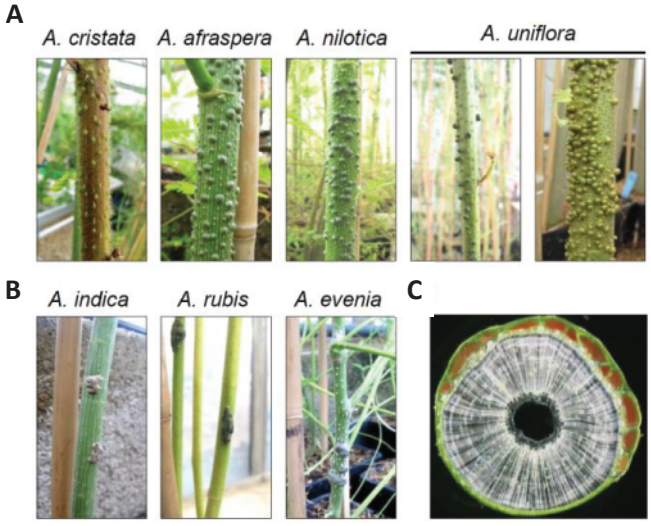


Figure 48. Nodulation caulinaire dans le genre *Aeschynomene* (thèse Albin Teulet, 2019).
(**A**) Photographies de nodules de tige de certains représentants du groupe 2 des *Aeschynomene.* (**B**) Photographies de nodules de tige de certains représentants du groupe 3 des *Aeschynomene.* (**C**) Coupe transversale de nodules de tige chez *A. sensitiva* infecté par *Bradyrhizobium* sp. ORS278. Le tissu infectieux est visible en rouge en raison de la présence de léghémoglobine.

l'Amérique centrale mais aussi principalement sur l'île de Madagascar, en Afrique et en Australie. Quelques espèces endémiques sont aussi retrouvées en Asie et en Australie.

Les *Aeschynomene* sont le plus souvent des arbustes et plantes herbacées et dans quelques cas peuvent être grimpantes. On les retrouve dans des forêts tropicales saisonnièrement sèches, dans des prairies boisées ou dans des zones rocheuses ou sablonneuses (figure 47). De nombreuses espèces sont semi-aquatiques[91] et on les retrouve ainsi dans les marais, au bord des rivières ou des trous d'eau ou dans des zones marécageuses et inondées (figure 47) (Lewis et al., 2005). Cette capacité à pousser dans l'eau implique que les *Aeschynomene* sont parfois présentes au milieu des rizières et sont souvent considérées comme des plantes adventices par les riziculteurs. Dans les pays tropicaux, les espèces du genre *Aeschynomene* sont utilisées pour améliorer les systèmes de pâturage telles qu'*A. americana*, *A. villosa* ou *A. evenia* (Lewis et al., 2005). Parfois elles sont aussi utilisées comme plantes ornementales comme *A. fluitans* retrouvée dans les bassins et aquariums. Enfin ce qui nous intéresse le plus ici, c'est que certaines *Aeschynomene* spp. sont capables d'établir une symbiose nodulaire uniquement avec des bactéries du genre *Bradyrhizobium* qui ont la particularité d'être photosynthétiques[92].

5.1.2. Des légumineuses à nodulation de tige

L'intérêt porté aux *Aeschynomene* a d'abord été lié à une caractéristique rare retrouvée chez peu d'espèces de légumineuses puisque seules une trentaine sur presque 20 000 la possède : la nodulation de tige. En effet, à l'inverse de la majorité des légumineuses nodulées par les rhizobiums, les espèces semi-aquatiques d'*Aeschynomene* présentent à la fois des nodules racinaires mais aussi des nodules caulinaires lorsqu'elles sont infectées par des *Bradyrhizobium* photosynthétiques (figure 48). La présence de primordia racinaires sur la tige qui s'activent au contact de l'eau ou dans des conditions de forte hygrométrie, permet le développement de racines adventives qui peuvent par la suite êtres colonisées et infectées par les bactéries.

Découverte initialement chez *Aeschynomene aspera* (Hagerup, 1928), cette particularité symbiotique de former des nodules caulinaires est également observée chez d'autres espèces semi-aquatiques appartenant au genre *Sesbania*, *Discolobium* et *Neptunia* (Boivin et al., 1997). C'est cependant au sein du genre *Aeschynomene* que la nodulation de tige est la plus représentée où 25 espèces ont été décrites à ce jour comme induisant des nodules caulinaires (Brottier et al., 2018). A partir des propriétés phénotypiques de différentes souches de *Bradyrhizobium* (caractère photosynthétique ou

[91] Dites aussi hydrophytes car elles vivent en partie immergées dans l'eau.

[92] Le caractère photosynthétique est peu commun chez les rhizobiums et n'a été ainsi observé que chez les *Bradyrhizobium* et les *Methylobacterium* (Jaftha et al., 2002).

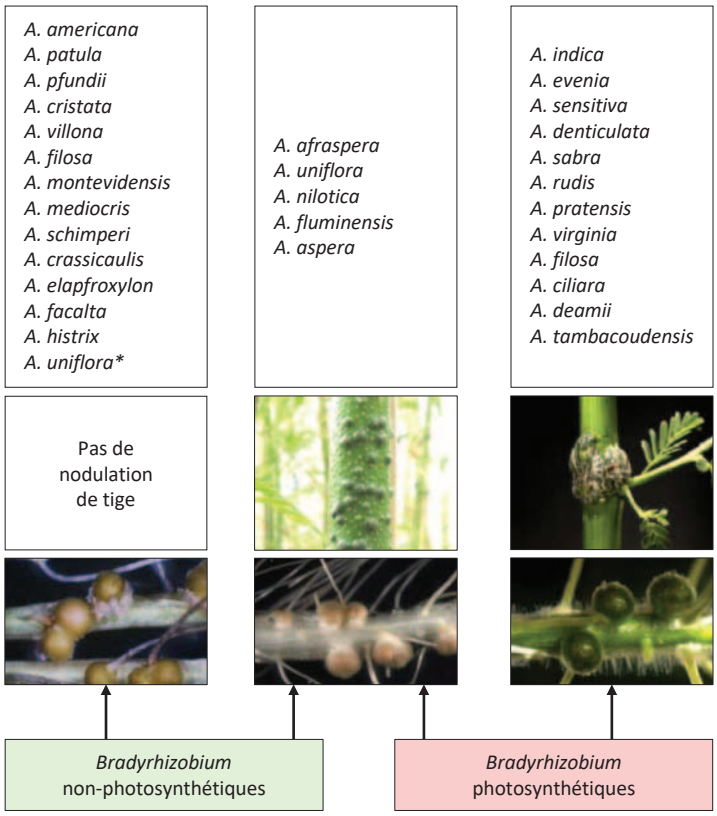


Figure 49. Groupes d'inoculation croisés entre *Bradyrhizobium* et *Aeschynomene* spp (thèse Albin Teulet, 2019).
Le groupe 1 des *Aeschynomene* est capable de former des nodules uniquement racinaires avec les *Bradyrhizobium* non photosynthétiques (**A. uniflora* est une exception puisqu'elle forme aussi des nodules de tige). Le groupe 2 peut quant à lui former des nodules à la fois sur la tige et les racines et est nodulé à la fois par des *Bradyrhizobium* non photosynthétiques et photosynthétiques (tous deux contenant des gènes *nod*). Le groupe 3 forme également des nodules sur les tiges et les racines mais est uniquement nodulé par des *Bradyrhizobium* photosynthétiques (ne contenant pas de gènes *nod,* à l'exception d'ORS285).

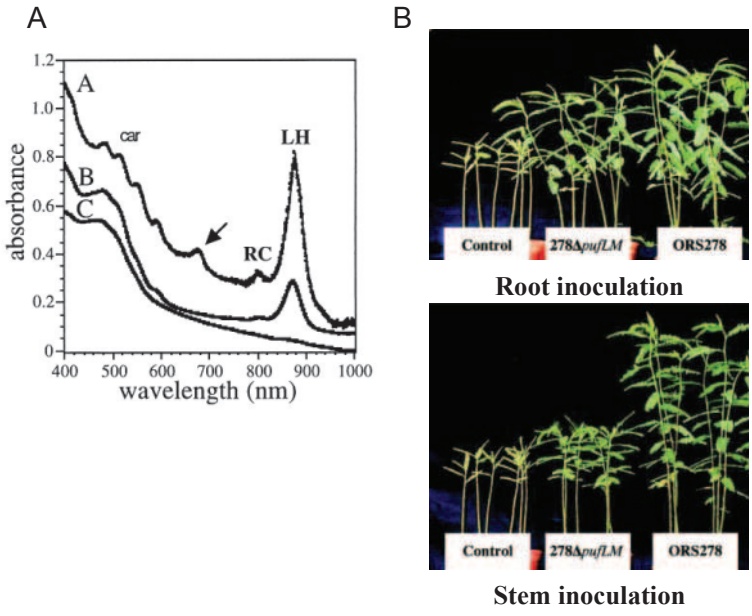


Figure 50. Activité photosynthétique de la souche *Bradyrhizobium* sp. ORS278 (Giraud et al., 2000).
(**A**) Spectre d'absorption dans le visible et le proche infrarouge des membrane de la souche ORS278 sauvage isolée de nodules de tige (A) ou cultivée sur boîte (B), et de la souche ORS278Δ*pufLM* mutée dans les gènes *pufL* et *pufM* codant le centre réactionnel du photosystème (C). La flèche indique un pic d'absorption correspondant à des chloroplastes issus de la plante. RC, reaction center, LH, light-harvesting complexes ; car, carotenoids. (**B**) Comparaison de la croissance des plantes d'*Aeschynomene sensitiva* inoculées au niveau de la tige ou au niveau de la racine, par la souche sauvage ORS278 ou ORS278Δ*pufLM* (5 semaines après inoculation).

non et spectre d'hôte), trois groupes d'inoculation croisées[93] (figure 49) ont été définis chez *Aeschynomene* spp. (Alazard, 1985; Chaintreuil et al., 2013). Le groupe 1, dont les représentants sont *Aeschynomene americana* et *Aeschynomene patula*, forme des nodules exclusivement racinaires suite à l'inoculation de *Bradyrhizobium* non-photosynthétiques. Les espèces du groupe 2 représentées par *Aeschynomene afraspera*, forment généralement une profusion de nodules de tiges et peuvent s'associer à la fois avec des *Bradyrhizobium* photosynthétiques et non-photosynthétiques. Enfin, le groupe 3 représenté par *Aeschynomene indica* et *Aeschynomene evenia*, développe de manière éparse des nodules de tiges avec uniquement des souches photosynthétiques. Cette dernière observation souligne la spécificité du groupe 3 qui ne peut être nodulé que par un nombre restreint de *Bradyrhizobium* photosynthétiques.

5.1.3. Les *Bradyrhizobium* photosynthétiques

Au sein du genre *Bradyrhizobium*, il a été identifié des bactéries montrant une activité photosynthétique en raison de leur capacité à synthétiser un photosystème qui contient la bactériochlorophylle, pigment chlorophyllien capable d'absorber l'énergie lumineuse et de la convertir par la suite en énergie chimique disponible. La première souche de *Bradyrhizobium* identifiée et possédant cette particularité a été isolée de nodules de tiges d'*A. indica* : *Bradyrhizobium* sp. BTAi1 (Evans et al., 1990).

L'anatomie de ces nodules de tige s'éloigne de celle retrouvée sur les nodules racinaires en raison notamment de leur épiderme exposé à la lumière et très riches en chloroplastes. Ces derniers contiennent de la chlorophylle végétale capable d'absorber principalement la lumière bleue (450 nm) et rouge (660 nm). A l'inverse, la bactériochlorophylle utilise la lumière rouge lointain (> 800 nm) finalement restée disponible par la plante (figure 50). Dans ce sens, les travaux d'Evans et ses collaborateurs (1990) ont montré que le photosystème bactérien peut fournir de l'énergie nécessaire à l'activité de la nitrogénase : des nodules de tiges illuminés par une longueur d'onde supérieure à 730 nm et ne pouvant pas être captée par les chloroplastes, stimule l'activité de la nitrogénase.

Ainsi, il a été suggéré que la photosynthèse bactérienne fournit de l'énergie pour le fonctionnement de la nitrogénase là où des bactéries non-photosynthétiques en sont incapables (Giraud & Fleischman, 2004). Un autre point remarquable chez ces *Bradyrhizobium* photosynthétiques est le mécanisme de régulation des gènes formant le photosystème qui fait intervenir un senseur de lumière (un bactériophytochrome) (Giraud et al., 2002). Celui-ci en fonction du ratio lumière rouge/infra-rouge, va activer ou pas leur expression. Grâce à ce contrôle, les bactéries peuvent

[93] Un groupe d'inoculation croisée est composé d'espèces de plantes qui ont la propriété d'être inoculées par un même groupe de bactéries.

mettre en place leur photosystème spécifiquement dans les nodules de tige et non dans les nodules racinaires qui demeurent à l'obscurité. Il semble donc qu'il y ait une adaptation à la fois du côté plante et du côté bactérien, pour optimiser la mise en place de la collecte de la lumière au niveau des nodules de tige pour alimenter la nitrogénase en énergie. La plante peut tirer ainsi un réel avantage de cette activité photosynthétique bactérienne dans les nodules de tige puisqu'elle peut économiser une partie de son énergie, normalement fournie à la bactérie pour le fonctionnement de la nitrogénase, et l'utiliser pour son propre développement (biomasse, graines).

La spécificité stricte de nodulation observée entre *Aeschynomene* spp. du groupe 3 et les *Bradyrhizobium* photosynthétiques a conduit également à supposer que ce caractère photosynthétique pouvait jouer un rôle direct dans la capacité de nodulation et d'infection. Dans ce sens, il a été montré que cette activité photosynthétique contribue à une meilleure nodulation des tiges sûrement en favorisant la survie de la bactérie exposée à la lumière ou en lui fournissant de l'énergie pour faciliter l'infection de la plante (Giraud et al., 2000). Néanmoins, aucun rôle du photosystème n'a été montré pour la nodulation racinaire. En effet, la mutation du photosystème de la souche *Bradyrhizobium* sp. ORS278 entraîne une diminution du nombre de nodules sur la tige sans pour autant affecter le nombre de nodules racinaires chez *Aeschynomene sensitiva* (Giraud et al., 2000).

Par ailleurs les *Bradyrhizobium* photosynthétiques qui sont les symbiotes naturels d'*Aeschynomene*, sont également capables de fixer l'azote à l'état libre au contraire d'une grande majorité de rhizobiums (Alazard, 1990). Cette capacité est à relier avec leur particularité de porter le gène *nifV* nécessaire à la synthèse de l'homocitrate synthase requise pour le fonctionnement de la nitrogénase (cf. 1.2.1). Ce groupe particulier de bactéries photosynthétiques est également retrouvé comme endophyte de certaines plantes non-légumineuses telle que le riz et confère un avantage à cette culture d'importance mondiale (C. Chaintreuil et al., 2000).

Le caractère qui demeure cependant le plus original chez ce groupe de bactéries photosynthétiques est leur capacité à induire des nodules de tiges et de racines chez certaines *Aeschynomene* alors que la plupart de ces bactéries, telles que *Bradyrhizobium* sp. ORS278 et BTAi1, sont dépourvues des gènes canoniques *nodABC*[94] nécessaires à la synthèse des FNs (Giraud et al., 2007). Cela implique donc que ces *Bradyrhizobium* sont capables d'utiliser un nouveau processus symbiotique alternatif indépendant des FNs, qui semble restreint au groupe d'inoculation croisée 3 (Giraud et al., 2007; Mornico et al., 2012). Cette découverte majeure a permis de mettre fin au dogme selon lequel toute interaction symbiotique légumineuse-rhizobium ne pouvait être initiée qu'en présence d'un signal bactérien médié par les FNs et par leur reconnaissance par la plante. Cette nouveauté dans la com-

[94] Exceptée la souche *Bradyrhizobium* sp. ORS285 qui contient à la fois les gènes *nod* et les gènes codant pour le T3SS.

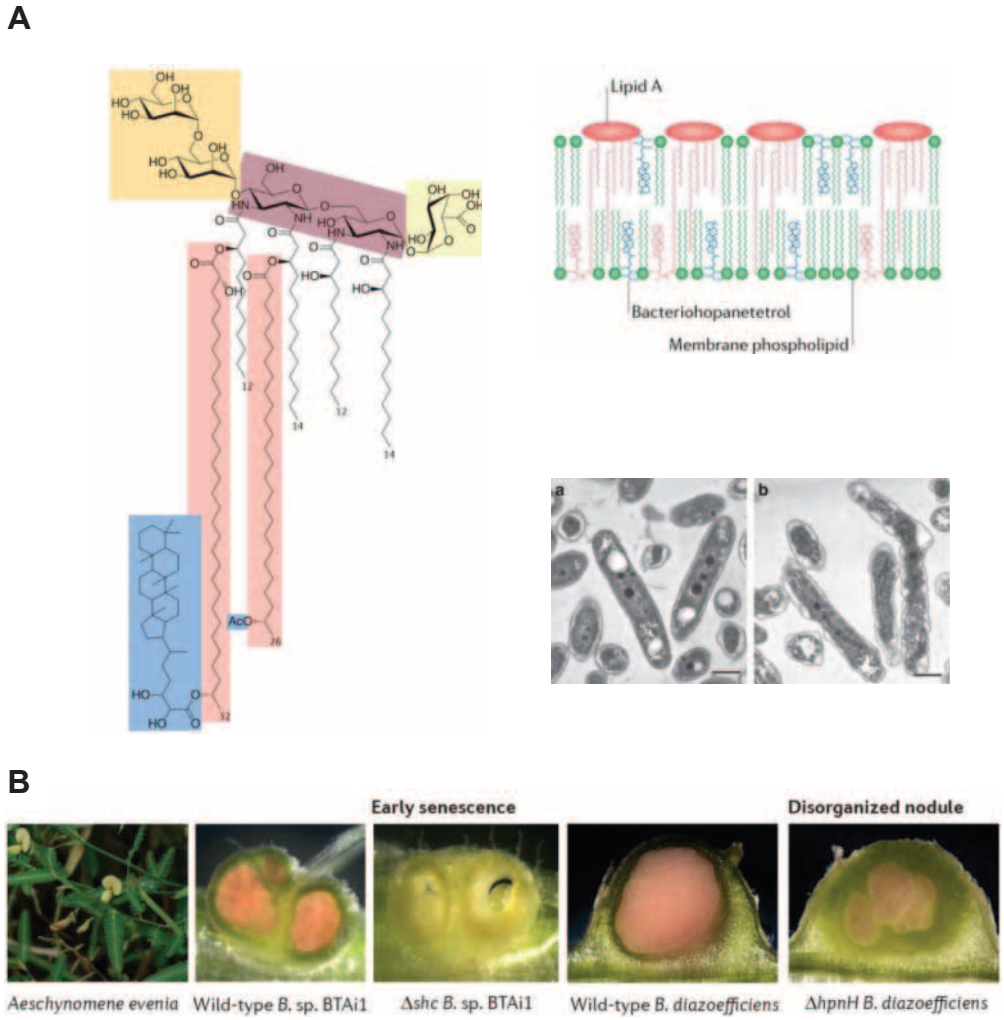


Figure 51. Structure des hopanoïdes-lipide A et leur rôle chez les *Bradyrhizobium* (adapté de Silipo et al., 2014; Kulkarmi et al., 2015; Belin et al., 2018).

(**A**) Structure d'un LPS contenant un hopanoïde (en bleu) lié de manière covalente aux chaînes d'acides gras (VLCFA, en rouge) du lipide A, un polysaccharide (en violet), un mannose (en orange) et un acide galacturonique (en jaune). Ces molécules confèrent une rigidité à la membrane représentée schématiquement à droite. La mutation dans le gène *shc* requis pour la biosynthèse des hopanoïdes altère cette structure membranaire (b) en comparaison de la souche sauvage BTAi1 (a) observée au microscope électronique à transmission. (**B**) La mutation dans le gène *shc* chez la souche BTAi1 altère également l'intégrité des bactéries au sein des nodules d'*A. evenia*, tandis que chez la souche *B. dizaoefficiens* USDA110, la mutation conduit chez *A. afraspera* à des nodules présentant un tissu central infecté désorganisé.

munauté scientifique a soulevé de nombreuses questions dont certaines commencent à être élucidées (cf. 5.3).

Comprendre le succès écologique des *Bradyrhizobium*

La membrane externe des bactéries Gram-négatives est composée majoritairement de lipopolysaccharides (LPS) qui sont essentiels à leur survie et à leur croissance dans des environnements variés. En effet, elle forme une barrière de défense face aux multiples stress (biotiques et abiotiques) auxquels la bactérie peut être confrontée et joue un rôle important dans l'interaction symbiotique avec les légumineuses. A la différence de la majorité des rhizobiums, les bactéries du genre *Bradyrhizobium* ont la particularité de synthétiser un LPS unique contenant un hopanoïde[1] attaché de manière covalente au lipide A qui forme la partie interne du LPS ancrée à la membrane phospholipidique (figure 51A). Cette structure hopanoïde-lipide A (HoLA) renforce la stabilité et la rigidité de la membrane externe et joue un rôle important dans la survie de la bactérie à la fois en vie libre ou symbiotique (Silipo et al., 2014; Kulkarni et al., 2015). En effet, la mutation d'un gène clé pour la synthèse des hopanoïdes chez la souche BTAi1 altère significativement l'intégrité membranaire de la bactérie et sa tolérance aux stress au sein du nodule mais aussi en vie libre (figure 51B). Cette composition lipidique de l'enveloppe bactérienne retrouvée chez les *Bradyrhizobium* serait à l'origine du succès écologique de ce genre bactérien capable de s'adapter plus facilement à des environnements variés souvent source de stress (Belin et al., 2018).
Une autre raison majeure du succès écologique des *Bradyrhizobium* est leur capacité à utiliser différents processus symbiotiques pour interagir avec leur plante hôte. A la fois capable de noduler des millions d'hectares de soja de manière Nod-dépendante mais aussi d'interagir avec des légumineuses tropicales du genre *Aeschynomene* de manière Nod-indépendante, les *Bradyrhizobium* ont su s'adapter au cours de l'évolution pour coloniser la planète entière pourtant soumis à des climats et à des écosystèmes variés. Leur association symbiotique avec les légumineuses cultivées est néanmoins au cœur de leur importance écologique bien que l'association des *Bradyrhizobium* avec des légumineuses non cultivées, en particulier au niveau des zones semi-arides aux sols pauvres, joue aussi un rôle écologique primordial pour l'apport d'azote à l'écosystème (Sprent, 2009).

[1] Les hopanoïdes sont une classe de lipides membranaires retrouvés dans la membrane plasmique des bactéries et assimilable au rôle et à la structure du cholestérol chez les eucaryotes.

5.2. Le mécanisme Nod-indépendant, T3SS-indépendant entre les *Bradyrhizobium* photosynthétiques et les *Aeschynomene* spp. du groupe 3

Il a pu être montré que la symbiose Nod-indépendante identifiée chez le groupe 3 d'inoculation croisée ne nécessite ni les gènes *nod*, ni ceux du T3SS. En effet, à l'exception de la souche ORS285 qui possède à la fois les gènes *nod* et ceux codant pour le T3SS, toutes les autres souches du supergroupe des *Bradyrhizobium* photosynthétiques identifiées et séquencées à ce jour ne possèdent aucun de ces gènes[95]. Par ailleurs, il a pu être montré que la mutation du T3SS ou des gènes *nod* chez la souche ORS285 n'affectait pas sa capacité à noduler les *Aeschynomene* du groupe 3 mais compromettait celle de noduler les *Aeschynomene* du groupe 2 (Giraud et al., 2007; Shin Okazaki et al.,

[95] A noter la présence d'un T3SS atypique chez la souche *B. oligotropicum* S58 et qui fait partie du clade des *Bradyrhizobium* photosynthétiques (Teulet et al., 2020).

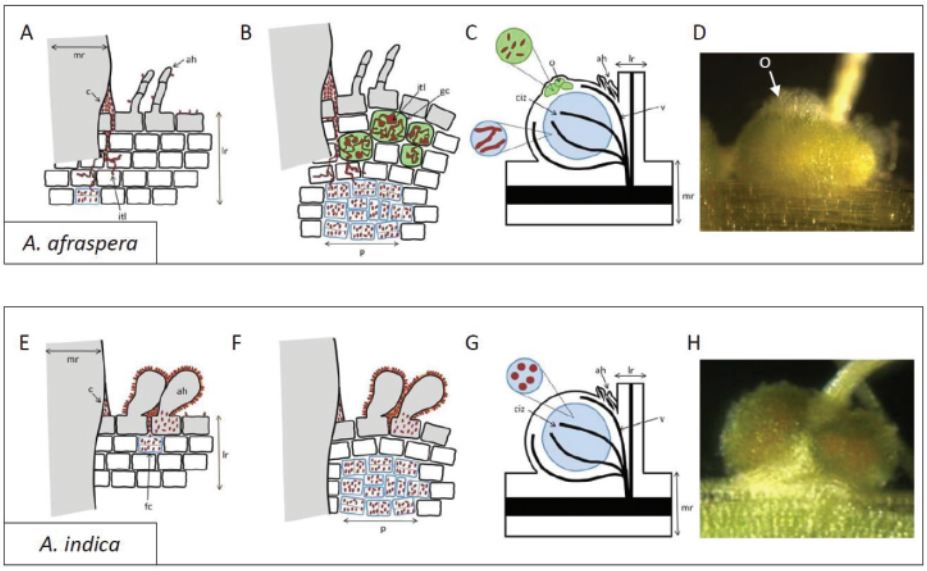


Figure 52. Processus d'infection et d'organogénèse nodulaire chez deux espèces d'*Aeschynomene* inoculées avec la souche ORS285 (adapté de Bonaldi et al., 2011).

Dans les deux cas, l'infection par la bactérie se fait par crack-entry mais la zone de colonisation n'est pas la même. Des divergences morphologiques peuvent être observées au niveau des poils axillaires des racines latérales d'*A. afraspera* (Nod-dépendant) et d'*A. indica* (Nod-indépendant) jouant un rôle sur l'attraction des bactéries représentées en rouge sur le schéma. La cellule initialement infectées se situe dans le cortex interne chez *A. afraspera* et dans le cortex externe chez *A. indica*. Après formation du nodule, une excroissance est observés chez ceux d'*A. afraspera* formant un réservoir de bactéries non différenciées contenues dans des cellules géantes (en vert). Enfin, la morphologie des bactéroïdes est allongée chez *A. afraspera* tandis qu'elle est sphérique chez *A. indica*. mr, racine principale ; lr, racine latérale ; c, crack ; ah, poil axillaire ; p, primordium ; fc, cellules fondatrices ; gc, cellules géantes ; o, excroissance nodulaire ; itl, tubulures similaires aux cordons d'infection ; v, vascularisation ; ciz, zone centrale infectée ; les cellules épidermiques sont en gris et les cellules corticales en blanc.

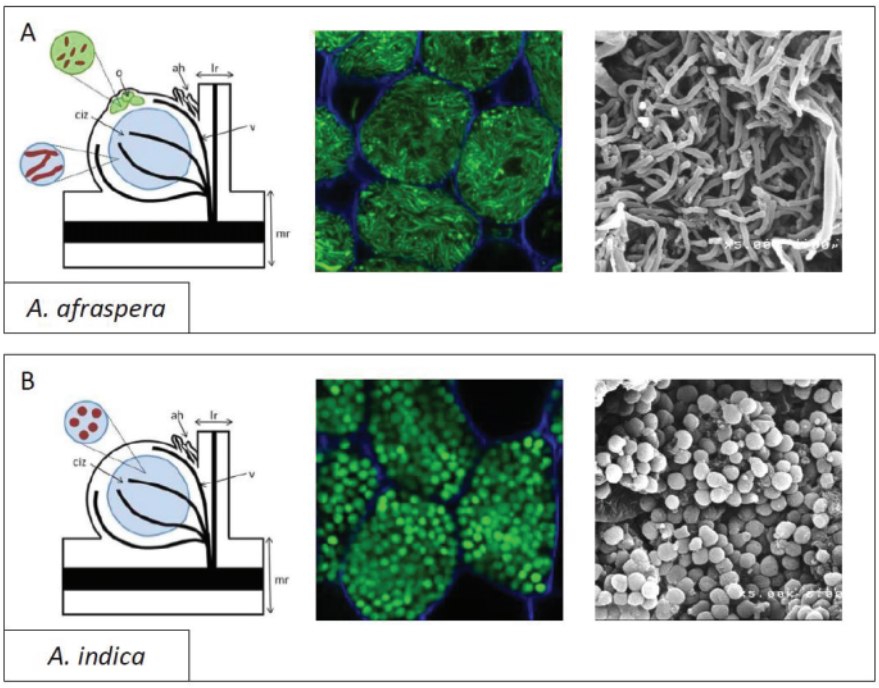


Figure 53. Différenciation en bactéroïdes chez deux espèces d'*Aeschynomene* inoculées avec la souche ORS285 (adapté de Bonaldi et al., 2011).

Chez les plantes du genre *Aeschynomene*, la différenciation en bactéroïdes est terminale. Cependant, on observe une morphologie des bactéroïdes variable selon les espèces : allongée chez *A. afraspera* et sphérique chez *A. indica*. Les observations ont été faites avec la souche ORS285 marquée à la GFP au microscope confocal (au centre) ou au microscope électronique à balayage (à droite).

2016). Cette souche est donc capable d'utiliser l'un ou l'autre des deux processus (avec ou sans FNs) selon le groupe d'*Aeschynomene* avec lequel elle interagit. L'ensemble de ces observations permet ainsi d'établir l'existence d'un processus symbiotique alternatif indépendant à la fois des FNs et du T3SS bactérien.

5.2.1. Un processus symbiotique aussi efficient que la symbiose Nod-dépendante

Dans les modèles symbiotiques "classiques", les FNs jouent un rôle majeur à la fois dans l'organogenèse du nodule mais aussi dans son infection. On peut donc s'interroger sur la façon dont la symbiose Nod-indépendante se met en place en l'absence de ce signal symbiotique. Pour éclairer ce point-là, la capacité de la souche photosynthétique ORS285 capable d'utiliser ces deux types de symbiose selon l'espèce d'*Aeschynomene* concernée, a été largement utilisée pour élucider cette question (Bonaldi et al., 2011). Comme décrit dans la partie 3.1.2, les *Aeschynomene* font partie des légumineuses infectées de manière intercellulaire par crack-entry. Les processus d'infection et d'organogenèse ont été surtout décrits pour les nodules racinaires formés chez *A. afraspera* (groupe 2 – nodulation Nod-dépendante) et *A. indica* (groupe 3 – nodulation Nod-indépendante) en utilisant la souche sauvage ORS285 et son mutant ORS285Δ*nodB* associé (figure 52) (Bonaldi et al., 2011). Il ne faut que 5 à 6 jours en moyenne après inoculation pour obtenir des nodules fonctionnels fixateurs d'azote, peu importe le processus de nodulation utilisé. L'infection chez *A. indica* se fait à partir d'une cellule du cortex externe alors que celle infectée de manière Nod-dépendante est plus interne. Une excroissance est observée à la surface des nodules chez *A. afraspera*, où les cellules végétales de plus grandes tailles concentrent une population bactérienne qui n'est pas différenciée en bactéroïdes et pourrait constituer un réservoir de bactéries capables d'infecter à nouveau la plante (Bonaldi et al., 2011).

Une fois l'organogenèse nodulaire achevée, il est observé des morphologies différentes des bactéroïdes : généralement sphériques chez les espèces Nod-indépendantes et allongées chez les espèces Nod-dépendantes (figure 53) (Czernic et al., 2015). De manière surprenante, dans les deux cas cette différenciation implique des protéines NCR-like synthétisées par la plante comme chez les légumineuses du clade IRLC dont fait partie *M. truncatula* (Czernic et al., 2015). L'implication de NCR-like conduit à une différenciation dite terminale des bactéroïdes, avec une augmentation du niveau d'endoréplication et une perte de la capacité à se multiplier après cette différenciation. Enfin peu importe la voie symbiotique utilisée (Nod-indépendante ou Nod-dépendante), une nette amélioration de la croissance des plantes est observée en raison d'une fixation d'azote efficiente par les bactéries. L'ensemble de ces observations nous permet de conclure que la symbiose Nod-indépendante mise en place par les *Bradyrhizobium* photosynthétiques est aussi efficiente et rapide que la sym-

biose Nod-dépendante. Dans les deux cas le processus d'infection et d'organogenèse nodulaire demeurent très similaires impliquant une infection initiale intercellulaire par "crack-entry" et une organogenèse nodulaire résultant d'une réactivation du cycle mitotique chez une ou quelques cellules fondatrices initialement infectées.

5.2.2. La recherche du signal "non Nod" : approche de mutagénèse bactérienne

En absence de FNs produits par la bactérie, on peut supposer qu'un autre signal bactérien puisse déclencher le processus d'organogenèse et d'infection conduisant à la formation de nodules fonctionnels fixateurs d'azote. Peu de temps après la découverte de ce processus symbiotique alternatif, il a été généré dans ce sens une banque de mutants Tn5[96] aléatoire chez la souche ORS278[97] pour essayer d'identifier les gènes bactériens impliqués dans la synthèse du signal dit "non Nod" (Bonaldi et al., 2010; Giraud et al., 2007). Ainsi, plus de 25 000 mutants bactériens ont été criblés sur *A. sensitiva* et *A. indica* dans le but d'identifier un mutant incapable de noduler la plante. Cependant, aucun des mutants n'a présenté un phénotype totalement déficient dans la nodulation bien qu'environ 300 soient apparus altérés dans le développement nodulaire ou dans l'activité fixatrice d'azote (Bonaldi et al., 2010; Giraud et al., 2007). Plusieurs hypothèses ont alors été émises quant à la raison de ce résultat : i) il existe une redondance fonctionnelle qui contrôle l'étape initiale de ce processus symbiotique, ii) le(s) gène(s) requis pour la synthèse du facteur "non Nod" sont indispensables à la survie de la bactérie ou iii) la procédure de mutagénèse utilisée n'a pas été saturante et la mutation attendue n'est pas apparue dans la banque de mutants (Giraud et al., 2007).

Malgré cela, cette mutagénèse bactérienne a permis d'obtenir de nouvelles connaissances sur l'établissement de la symbiose Nod-indépendante, T3SS-indépendante. Ainsi, plus de 50% des mutants altérés dans le processus de nodulation se sont avérés être déficients dans la biosynthèse des purines (Bonaldi et al., 2010). Ce résultat était d'autant plus intéressant puisque dans le cas des symbioses conventionnelles, des études précédentes suggéraient que le précurseur des purines pourraient agir également en tant que molécule signale durant la symbiose rhizobium-légumineuse (Newman et al., 1994; Noel et al., 1988). Dans ce sens l'adénosine monophosphate (AMP) qui est un dérivé de purine, est aussi un précurseur de la voie de synthèse des cytokinines qui ont été montrées comme jouant un rôle majeur dans l'organogenèse nodulaire (cf. 3.2.2). Il a ainsi été proposé initia-

[96] Mutagénèse par le transposon Tn5 d'origine bactérienne qui transfère des gènes dans le chromosome et permet ainsi d'interrompre ou de modifier la fonction du gène existant en provoquant une mutation.

[97] Le choix initial s'est porté sur la souche ORS278 en raison d'outils génétiques déjà disponibles chez cette bactérie, mais il a été aussi testé par la suite une banque de mutants Tn5 (10 000) chez la souche ORS285, à la fois sur *A. afraspera* et *A. indica*. Ce second criblage a permis de sélectionner des mutants Nod$^-$ stricts sur *A. afraspera* dans lesquels le Tn5 a été retrouvé inséré dans les gènes *nodABC*, mais aucun mutant Nod$^-$ n'a été obtenu sur *A. indica* lors de ce deuxième criblage (communication personnelle Eric Giraud).

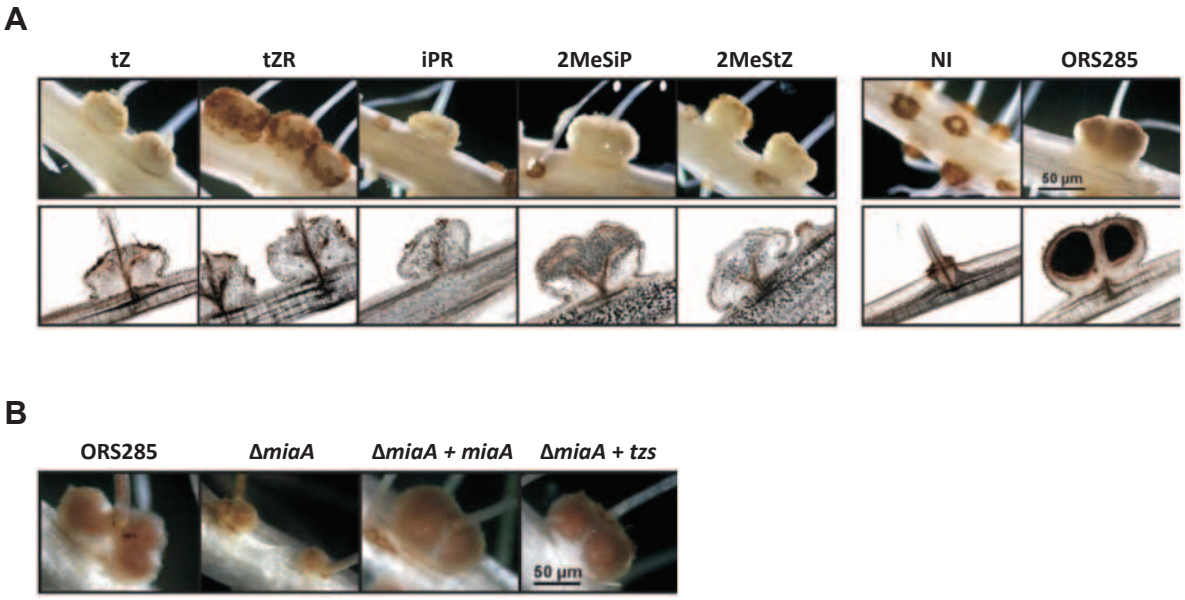


Figure 54. Les cytokinines contribuent à l'organogénèse nodulaire mais ne sont pas le signal "non Nod" (adapté de Podlešáková et al., 2007).
(**A**) L'application de diverses cytokinines exogènes induit la formation de nodules spontanés chez *A. indica* indiquant que ces molécules sont capables d'initier le programme de nodulation. (**B**) La mutation du gène *miaA* qui conduit à l'absence de biosynthèse des cytokinines chez la souche ORS285, affecte la développement nodulaire (nodules plus petits) mais ne l'abolit pas complètement. L'ajout du gène *miaA* ou *tzs* d'*Agrobacterium tumefaciens* restaure également le phénotype (tZ, *trans*-zeatin ; tZR, *trans*-zeatin riboside ; iPR, N6-isopentenyladenine riboside ; 2MeSiP, 2-methylthiol N6-isopentenyladenine ; 2MeStZ, 2-methylthiol *trans*-zeatin ; NI, non-inoculated).

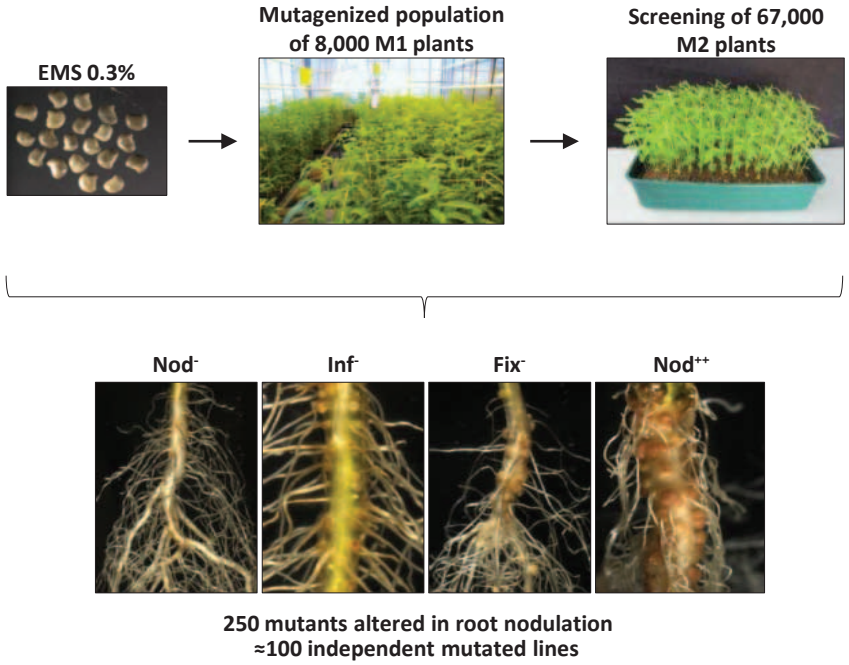


Figure 55. Approche génétique de mutagénèse EMS chez *A. evenia* et identification de mutants altérés dans le processus de nodulation (adapté de Quilbé et al., 2021).
L'analyse de presque 70 000 individus issus de la mutagénèse EMS a permis d'identifier 250 mutants présentant un phénotype symbiotique altéré. Cette collection de mutants constitue une ressource génétique considérable pour l'identification des déterminants symbiotiques Nod-indépendant. A ce jour, seuls les mutants Nod$^-$ ont été analysés.

lement que les cytokinines bactériennes pourraient être le signal reconnu par les *Aeschynomene* déclenchant le processus Nod-indépendant, divers résultats venant appuyer cette hypothèse comme l'apparition de pseudo-nodules après ajout exogène de différentes cytokinines chez *A. indica* ou la capacité des *Bradyrhizobium* photosynthétiques à sécréter des cytokinines (Podlešáková et al., 2013). Cependant la mutation chez ORS285 du gène *miaA* qui abolit totalement la biosynthèse de cytokinines induit seulement un retard de nodulation et une réduction de la taille des nodules sans pour autant réduire le nombre de nodules formés (Podlešáková et al., 2013). Les cytokinines produites par les *Bradyrhizobium* photosynthétiques jouent donc un rôle positif dans la nodulation mais ne sont pas le signal "non Nod" déclenchant la nodulation (figure 54).

Aujourd'hui, malgré d'autres approches de mutagénèses mais cette fois ciblées sur de multiples gènes candidats pouvant être impliqués durant l'interaction des souches de *Bradyrhizobium* ORS278 et ORS285 avec la plante hôte, on ne connaît toujours pas les mécanismes moléculaires gouvernant la symbiose Nod-indépendante. Ainsi, pour progresser dans la compréhension de ce processus symbiotique, il a été proposé de changer d'angle d'attaque en se focalisant sur le partenaire plante et en essayant d'identifier les récepteurs de ces signaux "non Nod".

5.2.3. La recherche des déterminants symbiotiques du côté plante : le modèle *Aeschynomene evenia*

Pour identifier les déterminants symbiotiques impliqués dans le processus Nod-indépendant, T3SS-indépendant, il a fallu choisir une plante modèle et développer des outils génétiques associés. Le choix s'est alors porté sur *A. evenia* dont les travaux pour déchiffrer les mécanismes symbiotiques sont menés depuis maintenant plus de 10 ans (Arrighi et al., 2012; Chaintreuil et al., 2016).

S'il a pu être montré par des approches RNAi[98] et de transcriptomiques que certains acteurs clés de la voie symbiotique commune tels que CCaMK, SYMRK ou encore NIN, sont également impliqués dans le processus Nod-indépendant (Fabre et al., 2015; Gully et al., 2018), pour identifier les déterminants spécifiques de cette symbiose, une approche de génétique directe a été utilisée.

Pour cela, une population de mutants EMS[99] d'*A. evenia* d'environ 70 000 individus a été criblée pour leur déficience symbiotique en utilisant la souche *Bradyrhizobium* sp. ORS278 (figure 55) (Quilbé et al., 2021). Il en est ressorti 250 mutants présentant une altération de la nodulation. Ces mutants de nodulation ont été classés en cinq catégories distinctes : i) les plantes pour lesquelles la nodulation est totalement abolie (Nod^-), ii) les plantes qui forment occasionnellement des nodules (Nod^{-*}), iii) les

[98] RNA interference, approche qui vise à éteindre ou réduire drastiquement l'expression d'un gène cible par la liaison de l'ARN introduit avec l'ARN messager du gène ciblé.

[99] Ethyl Methane Sulfonate ou méthanesulfonate d'éthyle en français. Composé organique mutagène qui produit des mutations aléatoires par substitution de nucléotides de paires de base G:C en A:T.

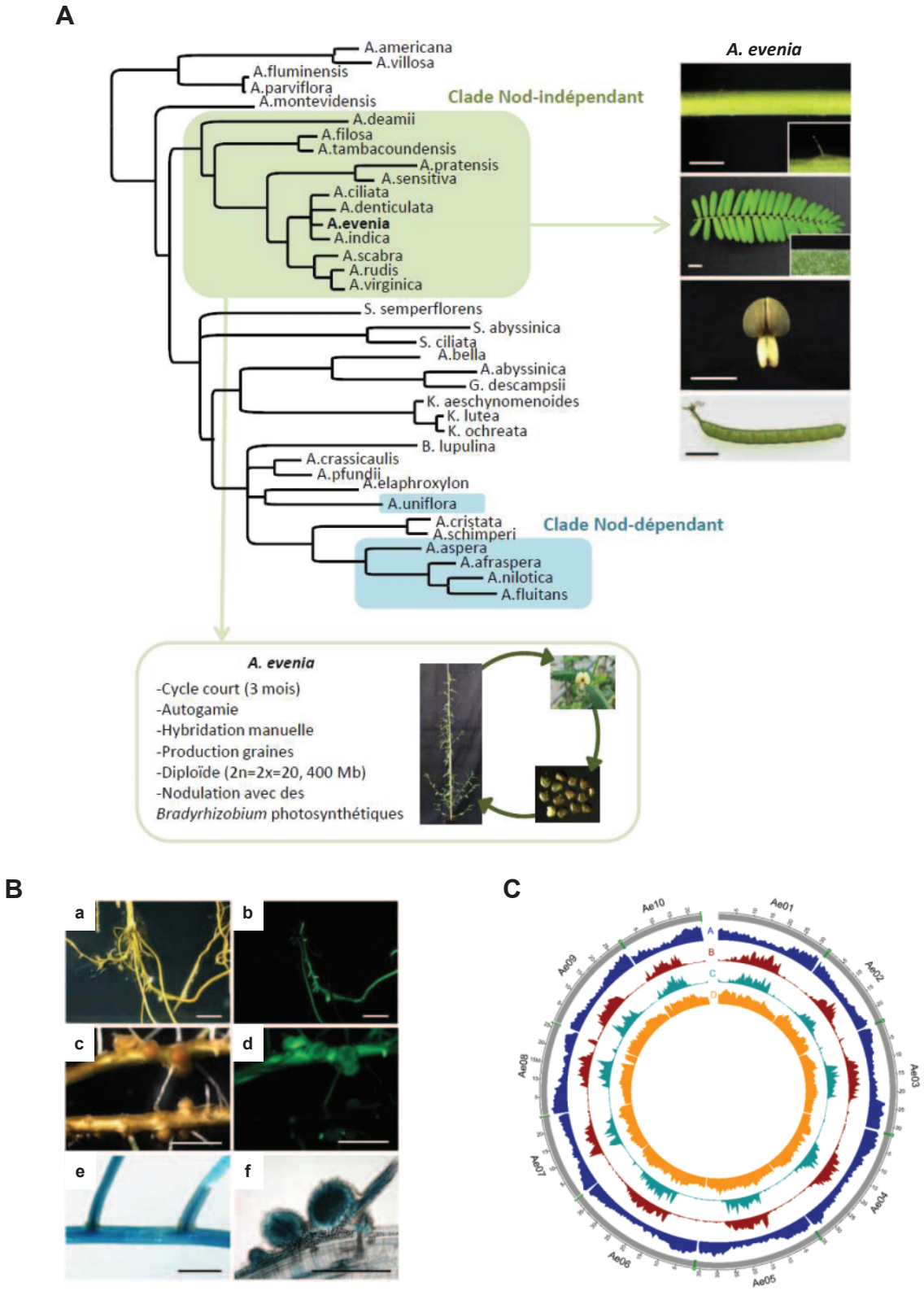


Figure 56. *Aeschynomene evenia*, une plante modèle pour la symbiose Nod-indépendante.
(**A**) Arbre phylogénétique basé sur les séquences ITS et *trnL* (gène chloroplastique) du genre *Aeschynomene* et d'espèces appartenant à des genres associés. L'encadré vert indique les espèces capables de noduler de manière indépendante des facteurs Nod, l'encadré bleu indique les espèces qui nodulent de manière Nod-dépendante uniquement. Parmi le clade des espèces Nod-indépendant, c'est *A. evenia* qui a été sélectionnée en 2012 en raison de ses caractéristiques biologiques, physiologiques et en raison de l'estimation de la taille relativement petite de son génome. (**B**) La transformation par *Agrobacterium rhizogenes* est efficiente et permet des approches de génétique fonctionnelle. Racines composites observées en lumière blanche (a) ou en lumière bleue pour l'épifluorescence de la GFP (b), 21 jours après transformation exprimant la construction *pCaMV35S::GFP*. Les racines transformées développent des nodules après inoculation avec ORS278 visibles en lumière blanche (c) et en épifluorescence (d). Pareillement, un système rapporteur du gène *gusA* (*pCaMV35S::gusA*) est fonctionnel après transformation et inoculation des racines (e-f). (**C**) Le génome d'*A. evenia* a été entièrement séquencé (https://aeschynomenebase.fr/) (adapté de Arrighi et al., 2012; Quilbé et al., 2021 et thèse de Johan Quilbé, 2021).

plantes formant des nodules non infectés (Inf$^-$), iv) celles possédant des nodules non fixateurs (Fix$^-$) et v) les plantes hypernodulantes (Nod^{++}). Cette collection de mutants représente donc une ressource génétique importante pour déchiffrer les mécanismes de la symbiose Nod-indépendante, T3SS-indépendante.

Pourquoi *Aeschynomene evenia* est-elle une bonne plante modèle ?

Un organisme modèle doit répondre à certains critères pour permettre l'étude approfondie de phénomènes biologiques particuliers. Pour cela, l'espèce en question doit être facilement manipulable en laboratoire et avoir une facilité de culture d'un grand nombre d'individus à la fois. Sa taille doit donc être relativement petite, son cycle de vie et de reproduction courts associés à une taille raisonnable de son génome. Sa génétique doit suivre les lois de Mendel et permettre le développement de mutations. Par ailleurs la diploïdie de son génome est un critère qui facilite les approches génétiques. Pour étudier la symbiose Nod-indépendante, *A. evenia* remplit l'ensemble de ces critères (figure 56) :

1. Elle appartient au groupe 3 d'inoculation croisée et est donc strictement nodulée de manière Nod-indépendante. De plus, elle fait partie des espèces pour laquelle le plus grand nombre d'accessions est disponible et possède une grande diversité génétique.
2. Elle possède un génome diploïde séquencé (2n=2x=20) d'une taille correcte (376 Mb).
3. C'est une plante autogame permettant l'obtention de lignées homozygotes et il est également possible de réaliser des hybridations manuelles.
4. Son cycle de vie est court (3 mois) avec un rendement en graines élevé et une taille raisonnable pour la culture *in vitro* et en serre.
5. Des outils de transformation génétique sont disponibles (*Agrobacterium rhizogenes*) et d'autres sont en court de développement (édition du génome par CRISPR-Cas9, transformation stable).

5.2.4. Un complexe de récepteur reconnaissant le signal "non Nod" ?

Pour identifier les acteurs symbiotiques majeurs impliqués dans le processus de nodulation, les mutants de la catégorie phénotypique Nod$^-$ ont d'abord été analysés. Les analyses faites sur ces mutants ont permis d'identifier cinq gènes de la voie symbiotique commune : *AePOLLUX*, *AeCCaMK*, *AeCYCLOPS*, *AeNSP2* et *AeNIN*. Cette voie semble donc être partiellement conservée chez *A. evenia* et vient confirmer en partie les résultats précédemment obtenus (Fabre et al., 2015; Gully et al., 2018).

Par ailleurs, aucun mutant n'a été identifié dans les gènes codant les récepteurs des FNs, pourtant présents chez *A. evenia* (Quilbé et al., 2021), confirmant la spécificité de ce nouveau processus alternatif qui est bien Nod-indépendant. En effet bien que présents dans le génome, AeNFR1 n'est pas exprimé (Gully et al., 2018) et AeNFR5 présente une mutation dans le motif requis pour la reconnaissance des ligands (communication personnelle Jean-François Arrighi).

De manière intéressante, l'analyse des mutants Nod$^-$ a aussi révélée des mutations présentes sur des gènes inconnus et qui seraient donc de bons candidats pour découvrir de nouveaux gènes symbiotiques associées à la nodulation Nod-indépendante. Parmi les locus identifiés, un gène présente des

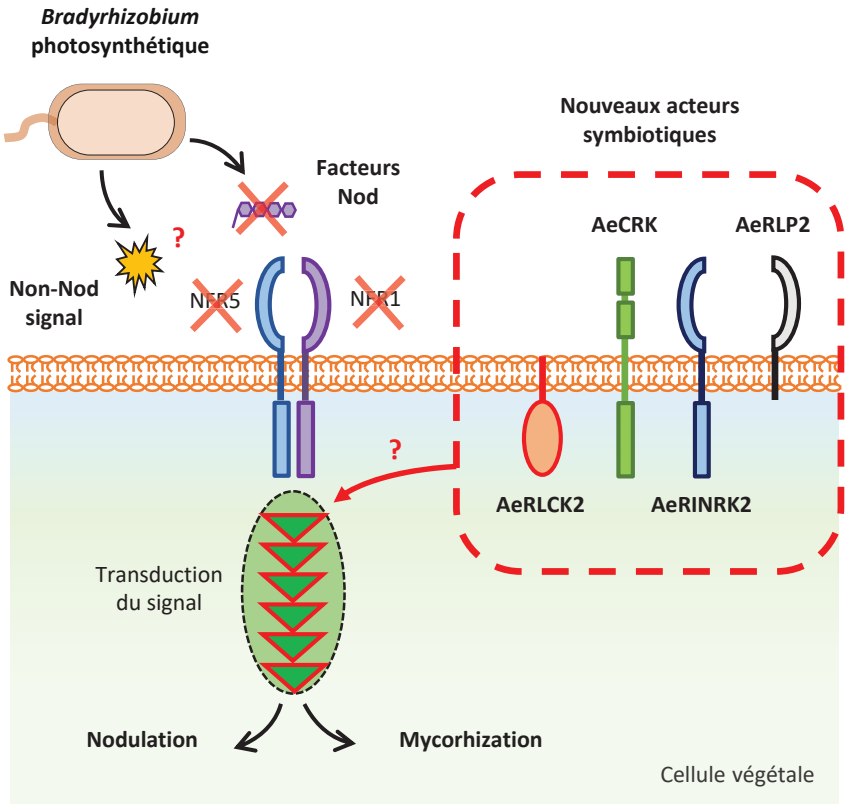


Figure 57. Modèle putatif d'un complexe de récepteurs "non Nod" (communication de Natasha Horta-Araújo).
L'ensemble des nouveaux gènes symbiotiques identifiés suite au crible EMS des mutants Nod$^-$ codent pour des protéines transmembranaires potentiellement impliquées dans des voies de signalisation. L'hypothèse que ces protéines agiraient en complexe de récepteurs pour reconnaître un signal "non Nod" et activeraient la voie de signalisation symbiotique est actuellement étudiée.

mutations et code pour un récepteur-like kinase riche en cystéine : CRK[100]. Aucun orthologue d'*AeCRK* chez les légumineuses modèles n'a pu être identifié, appuyant l'idée que ce gène est spécifique de la symbiose Nod-indépendante (Quilbé et al., 2021). La prédiction de domaine extracellulaire DUF26[101], d'un domaine transmembranaire et d'une sérine/thréonine kinase intracellulaire suggère que ce récepteur pourrait reconnaître un signal extracellulaire tout comme NFR1/NFR5 le font pour les FNs.

Trois autres nouveaux gènes ont depuis été identifiés chez les mutants Nod$^-$ (figure 57) : *AeRLCK2*[102], *AeRINRK*[103] et *AeRLP*[104] (communication personnelle Jean-François Arrighi). Tout comme *AeCRK*, ces nouveaux gènes appartiennent à des familles protéiques intervenant dans les étapes précoces de diverses voies de signalisation et codent pour des récepteurs membranaires (Liang & Zhang, 2022). L'activation des domaines kinases permettrait la transmission du signal produit par la reconnaissance d'un ligand extracellulaire.

Une des hypothèses étudiées actuellement est que ces nouveaux récepteurs formeraient un complexe protéique membranaire, impliqué dans la perception d'un signal synthétisé par les *Bradyrhizobium* photosynthétiques mais aussi dans l'activation de la voie de signalisation Nod-dépendante (figure 57). La caractérisation fonctionnelle de ces nouveaux récepteurs ainsi que l'étude de leur histoire évolutive (projet ANR "SymWay") sont actuellement en cours et font l'objet d'une thèse réalisée par Natasha Horta-Araújo. La recherche des interactants moléculaires de ces nouveaux récepteurs permettra éventuellement d'identifier le(s) signal "non Nod". A noter que jusqu'à présent, seule la catégorie des mutants EMS Nod$^-$ a été principalement analysée. On peut donc imaginer qu'au sein des autres mutants de nodulation, d'autres acteurs connus et nouveaux vont émerger et ainsi nous permettre de mieux comprendre les mécanismes sous-jacents de la symbiose Nod-indépendante, T3SS-indépendante.

5.2.5. Divers scénarios évolutifs de la symbiose Nod-indépendante

Comme évoquée dans la partie 3.1.2, l'infection intercellulaire serait un mécanisme supposé ancestral par rapport à l'infection intracellulaire par cordon d'infection. Dès lors, une symbiose Nod-indépendante utilisant un mode d'infection intercellulaire serait elle aussi ancestrale par rapport à la symbiose Nod-dépendante. Cependant d'autres scénarios évolutifs ont été proposé et aucun ne peut

[100] Cystein-rich Receptor-like Kinase.
[101] Domain of Unknown Function 26. Domaine retrouvé spécifiquement chez les végétaux mais des analyses structurales ont révélé qu'il existe une forte similarité avec les lectines fongiques. Ce domaine pourrait donc constituer un groupe de protéines capable de lier les carbohydrates (Vaattovaara et al., 2019).
[102] Receptor-like Cytoplasmic Kinase.
[103] Rhizobial Infection Receptor-like Kinase.
[104] Receptor Like Protein.

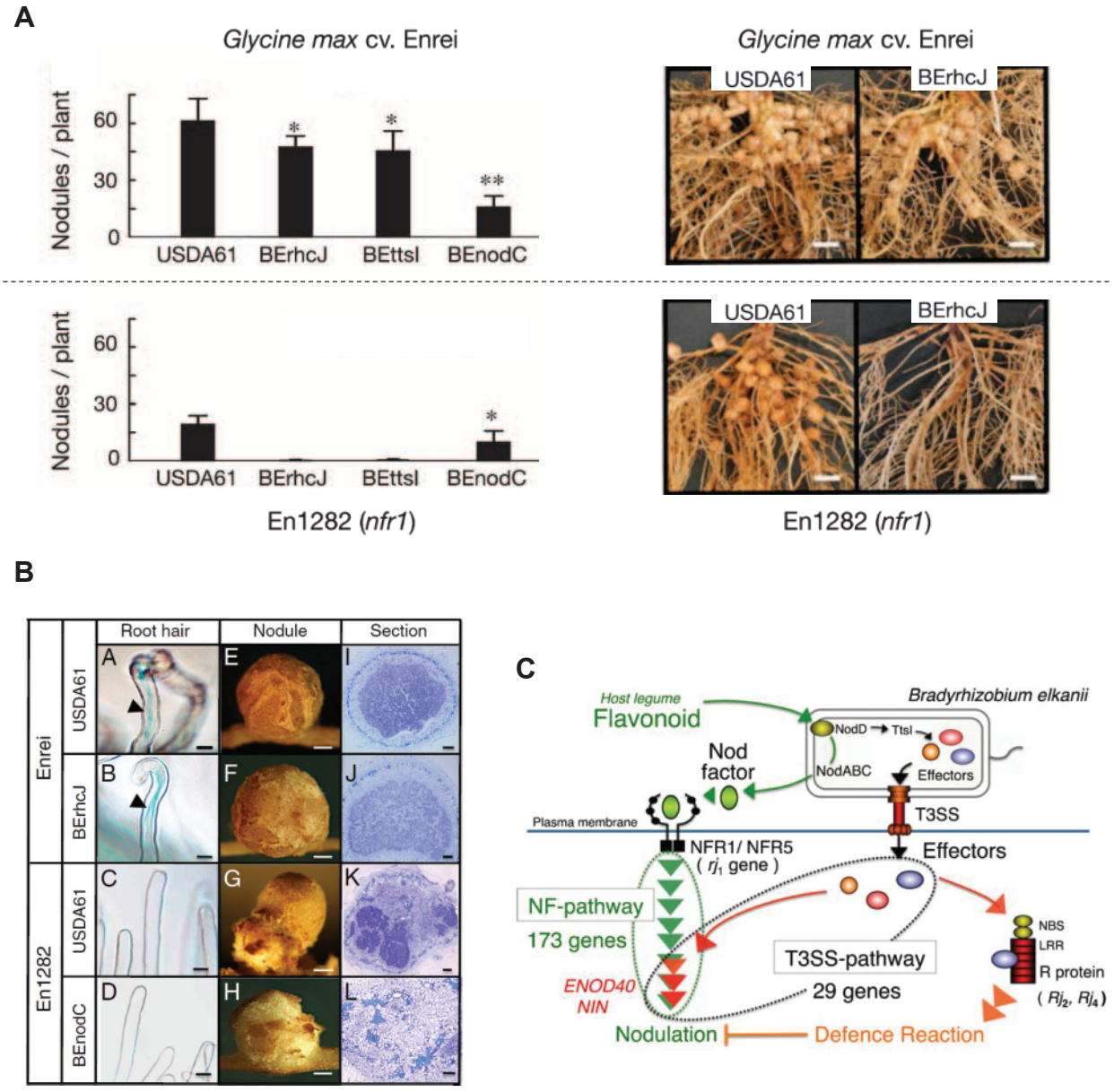


Figure 58. Piratage de la voie de signalisation symbiotique par le T3SS des *Bradyrhizobium* non photosynthétiques (adapté d'Okazaki et al., 2013).

(**A**) *Bradyrhizobium elkanii* USDA61 est capable de noduler le cultivar de soja Enrei (En1282) muté dans le récepteur NFR1 requis pour la reconnaissance des facteurs Nod. De plus, le mutant bactérien dans le gène *nodC* (BEnodC) est toujours capable de noduler le cultivar Enrei sauvage tout comme le mutant *nfr1*, soulignant que les facteurs Nod peuvent être dispensables durant l'interaction. En revanche, la mutation du T3SS (BErhcJ) ou de son régulateur transcriptionnel (BEttsI) abolit la nodulation chez le mutant plante *nfr1* (En1282), mettant en évidence le rôle clé du T3SS durant l'interaction. (**B**) Dans ce cas particulier, les nodules formés ne sont pas infectés par cordon d'infection mais de manière intercellulaire chez En1282 (C-D) en comparaison du cultivar sauvage Enrei (A-B). Cependant, les nodules formés dans un contexte T3SS-dépendant montre une désorganisation du tissus infectieux (K-L) par rapport au cultivar sauvage (I-J). (**C**) Modèle de la voie de signalisation symbiotique empruntée par le processus alternatif Nod-indépendant, T3SS-dépendant qui court-circuite la signalisation médiée par la reconnaissance des facteurs Nod chez le soja. Ce modèle suggère que des effecteurs de type 3 seraient directement impliqués dans le processus symbiotique sans pour autant être reconnus par les mécanismes de défense de la plante hôte tels que les protéines de résistance Rj2 et Rj4.

vraiment être réfuté ou accepté dans son entièreté. Ainsi, trois scénarios ont été suggérés (Madsen et al., 2010; Okubo et al., 2012) : i) une acquisition plus tardive de la symbiose Nod-dépendante en raison notamment de la capacité de certains mutants du lotier, affectés dans des gènes symbiotiques, à être nodulés et infectés à de très faibles fréquences par un mutant dans les gènes *nod* de *Mesorhizobium loti*, suggérant que des vestiges du processus Nod-indépendant pourraient avoir été conservés chez certaines légumineuses, ii) une acquisition ancestrale de la symbiose Nod-dépendante conduisant au développement plus récent chez un groupe restreint de plantes du processus Nod-indépendant et qui est basée sur l'idée d'une coévolution forte entre les *Aeschynomene* du groupe 3 et les *Bradyrhizobium* photosynthétiques, le fait que les *Aeschynomene* du groupe 3 forment un groupe monophylétique supporte un tel scénario, et enfin iii) une acquisition indépendante de la capacité symbiotique par l'ancêtre des *Aeschynomene* du groupe 3 qui serait une légumineuse initialement non-nodulante.

Dans l'hypothèse que la symbiose Nod-indépendante soit ancestrale, cela signifierait que les *Bradyrhizobium* photosynthétiques seraient les symbiotes ancestraux des légumineuses (Ormeño-Orrillo & Martínez-Romero, 2019).

5.3. Découverte d'un second processus symbiotique Nod-indépendant

5.3.1. Le processus Nod-indépendant, T3SS-dépendant chez le soja et les *Aeschynomene*

Un autre exemple de symbiose Nod-indépendante a pu être identifié chez le soja qui est pourtant une espèce de légumineuse connue requérant les FNs pour déclencher la symbiose (Okazaki et al., 2013). En effet, il a été mis en évidence par le groupe de Shin Okazaki que la souche non-photosynthétique *Bradyrhizobium elkanii* USDA61 mutée dans le gène *nodC* requis pour la synthèse des FNs, conserve sa capacité à former des nodules sur le cultivar sauvage *Glycine max* cv. Enrei mais aussi sur un mutant *nfr1* (cultivar En1282) incapable de reconnaître les FNs (figure 58A). En revanche, la mutation du T3SS (*rhcJ*) ou de son régulateur transcriptionnel (*ttsI*) chez la bactérie abolit totalement l'interaction symbiotique (Okazaki et al., 2013). Dans ce système artificiel où le soja n'est plus capable de reconnaître les FNs, l'infection ne passe pas par la formation d'un cordon d'infection comme cela se produit dans la nature, mais se fait de manière intercellulaire sûrement par crack-entry (figure 58B).

De manière intéressante, la souche non-photosynthétique USDA61 est aussi capable de noduler dans des conditions *in vitro A. indica*. Elle induit un nombre similaire de nodules que la souche modèle ORS278 bien que ceux-ci soient non fonctionnels car non infectés intracellulairement et souvent associés à des nécroses (Okazaki et al., 2016). Il a ainsi pu être confirmé comme chez le soja que le mu-

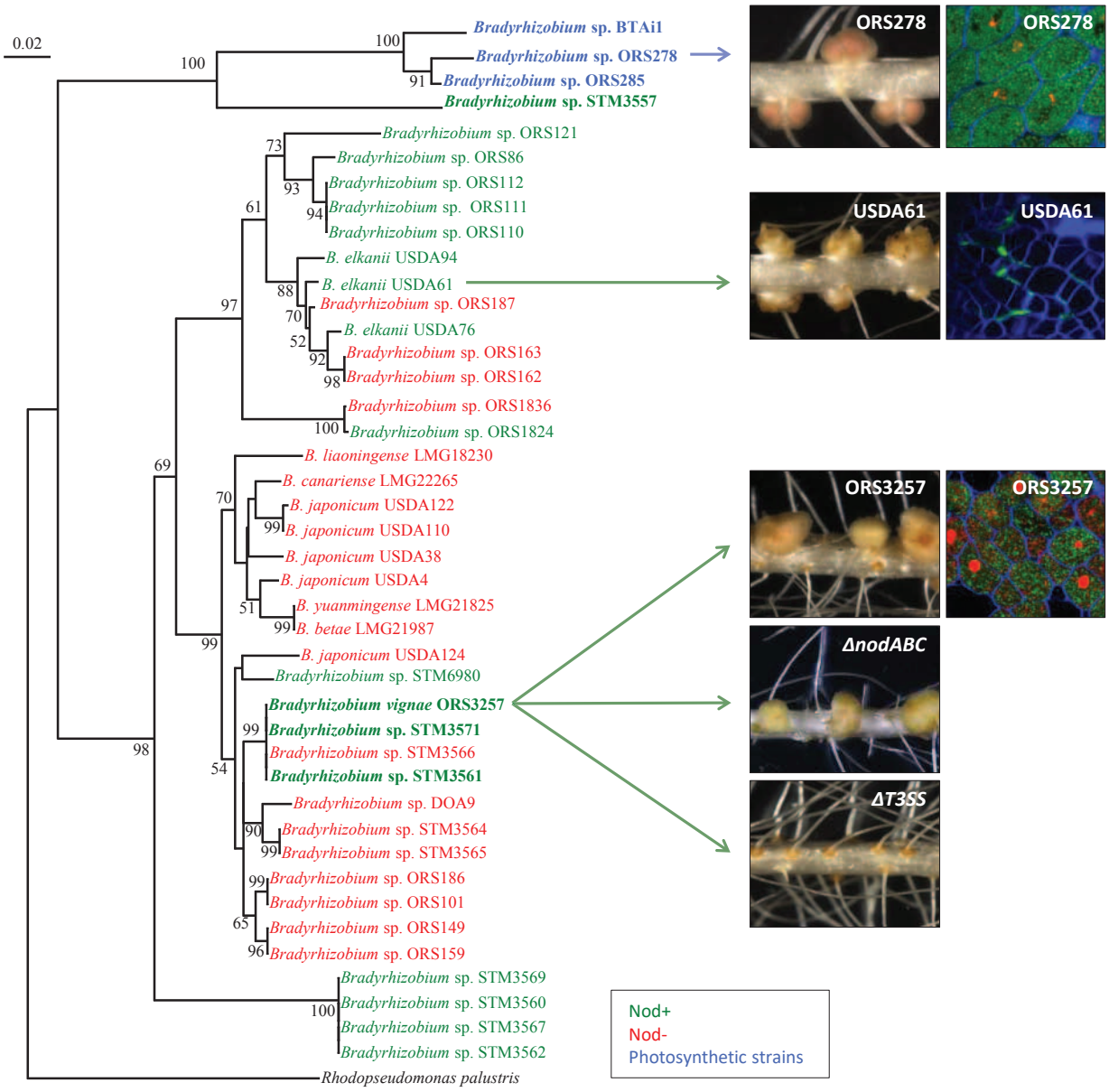


Figure 59. Mise en évidence de deux mécanismes de nodulation indépendants des facteurs Nod chez *A. indica*, dépendant ou pas d'un T3SS (adapté d'Okazaki et al., 2016).

La capacité à induire des nodules de manière Nod-indépendante chez *A. indica* est un trait répandu chez les *Bradyrhizobium*. L'arbre phylogénique construit, basé sur l'alignement de la séquence intergénique 16S-23S de l'ADN ribosomique, représente une collection de *Bradyrhizobium* non-photosynthétiques et photosynthétiques. Le code couleur est indiqué dans la légende en bas à droite. En gras apparaissent les souches pour lesquelles l'infection du nodule est intracellulaire. Les tests symbiotiques ont fait ressortir la souche *Bradyrhizobium vignae* ORS3257 comme modèle pour l'étude de la symbiose Nod-indépendante, T3SS-dépendante chez *Aeschynomene*, capable d'induire des nodules infectés intracellulairement à l'inverse de *B. elkanii* USDA61.

tant *rhcJ* d'USDA61 n'est plus capable de noduler *A. indica* alors que le mutant *nodC* l'est toujours, confirmant aussi que cette symbiose est Nod-indépendante, T3SS-dépendante (Okazaki et al., 2016).

Pour mieux comprendre cette habilité des souches non-photosynthétiques à noduler *A. indica* de manière Nod-indépendante, T3SS-dépendante, une collection d'environ 40 souches de *Bradyrhizobium* a été testée sur cette même plante (Okazaki et al., 2016). Il apparaît que plus de 40% des souches non-photosynthétiques sont capables d'induire des nodules bien que ces derniers soient rarement infectés intracellulairement (figure 59). Plus surprenant encore, on peut également souligner que dans cette étude, toutes les souches possèdent les gènes codant pour le T3SS soulignant le fait que ce second processus alternatif est répandu chez les *Bradyrhizobium* (Okazaki et al., 2016). Parmi les souches capables d'aller le plus loin dans le processus de nodulation, il est ressorti *Bradyrhizobium vignae* ORS3257 (anciennement nommée STM6978) pour laquelle une faible activité nitrogénase est détectée en test ARA[105] (Okazaki et al., 2016). Cependant l'étape de différenciation en bactéroïdes est altérée, ce qui expliquerait pourquoi aucun bénéfice réel n'est observé chez *A. indica* à la différence des souches photosynthétiques (figure 59). Enfin, cette étude réalisée sur une diversité de *Bradyrhizobium* a montré que la capacité à noduler de manière T3SS-dépendante n'est pas restreinte à *A. indica* puisque huit autres espèces du groupe 3 peuvent être nodulées à la fois par USDA61 et ORS3257. Cependant dans tous les cas, c'est la souche ORS3257 qui va le plus loin dans le processus de nodulation et d'infection. Par ailleurs, la présence d'un T3SS fonctionnel ne garantit pas le succès symbiotique puisque des souches non-photosynthétiques telles que *Bradyrhizobium diazoefficiens* USDA110 et *Bradyrhizobium* sp. DOA9 sont incapables de noduler *A. indica* (Okazaki et al., 2016; Songwattana et al., 2017).

Ainsi il a été montré pour la première fois qu'un T3SS fonctionnel présent chez la bactérie est requis pour déclencher la nodulation dans ce contexte particulier. Sachant que la plupart des *Bradyrhizobium* photosynthétiques nodulant les *Aeschynomene* ne possèdent ni les gènes *nod*, ni les gènes codant pour le T3SS, ce résultat indique donc l'existence d'au moins deux processus symbiotiques indépendants des FNs, nécessitant ou pas le T3SS.

Pendant longtemps il était admis que le rôle principal du T3SS dans la symbiose rhizobium-légumineuse était de moduler la réponse immune de la plante et de promouvoir ou pas l'interaction symbiotique entre les deux partenaires en favorisant initialement l'infection (Marie et al., 2001; Staehelin & Krishnan, 2015). Cependant, l'ensemble de ces résultats obtenus par Shin Okazaki suggère que le T3SS peut permettre aussi l'activation de la voie de signalisation symbiotique en contournant la reconnaissance des FNs (figure 58C). Suite à cette découverte, l'identification des effec-

[105] Pour Acetylene Reduction Assay. Technique permettant de mesurer indirectement l'activité nitrogénase grâce à sa capacité à couper la triple liaison de l'acétylène (C_2H_2) et de former de l'éthylène (C_2H_4).

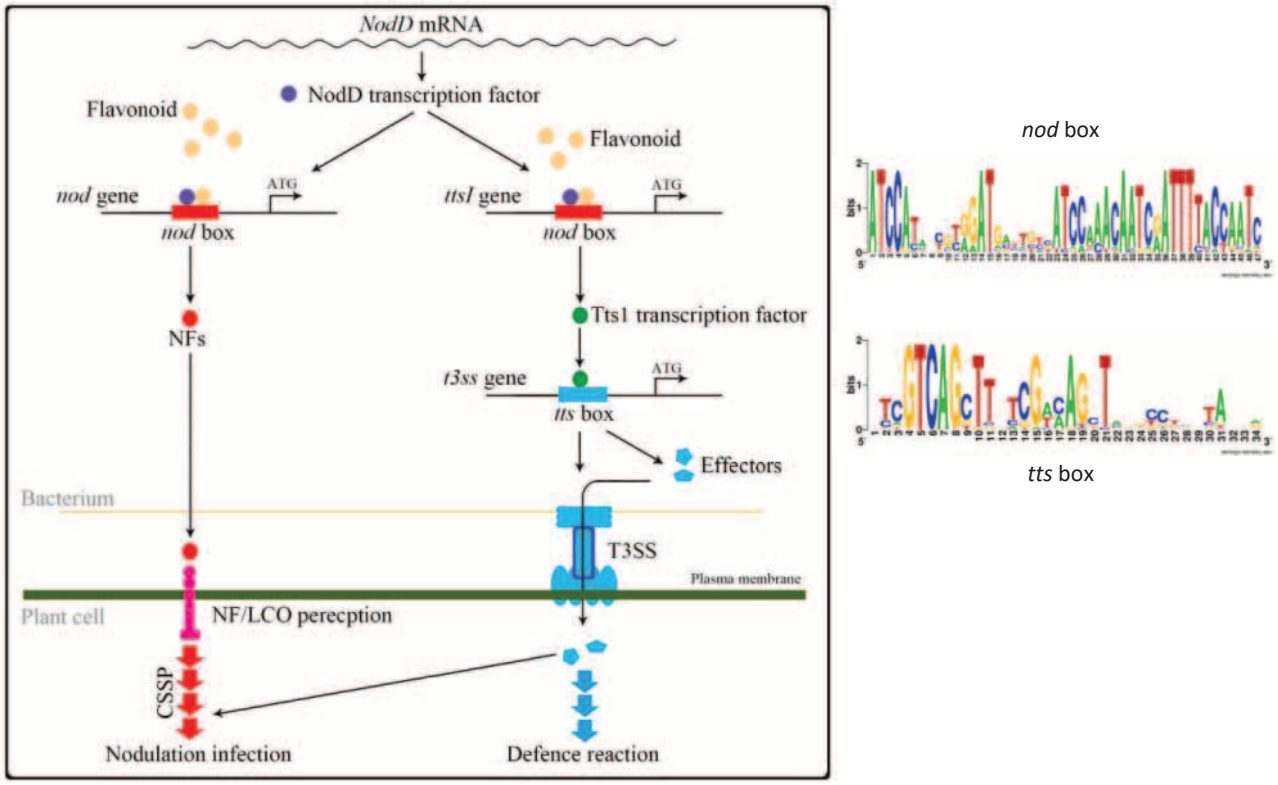


Figure 60. Mécanismes de régulation des gènes *nod* et ceux codant pour le T3SS par le facteur de transcription NodD (adapté de Dong & Song, 2020).
Les flavonoïdes sécrétés par la plante sont reconnus par le facteur de transcription bactérien NodD. Cette reconnaissance conduit à l'activation des gènes sont le contrôle d'une *nod* box souvent retrouvée devant les gènes *nod* mais aussi devant celui de *ttsI*. Plus tard, la synthèse du facteur de transcription TtsI va permettre l'expression des gènes codant le T3SS à travers sa liaison sur une *tts* box située dans la région promotrice.

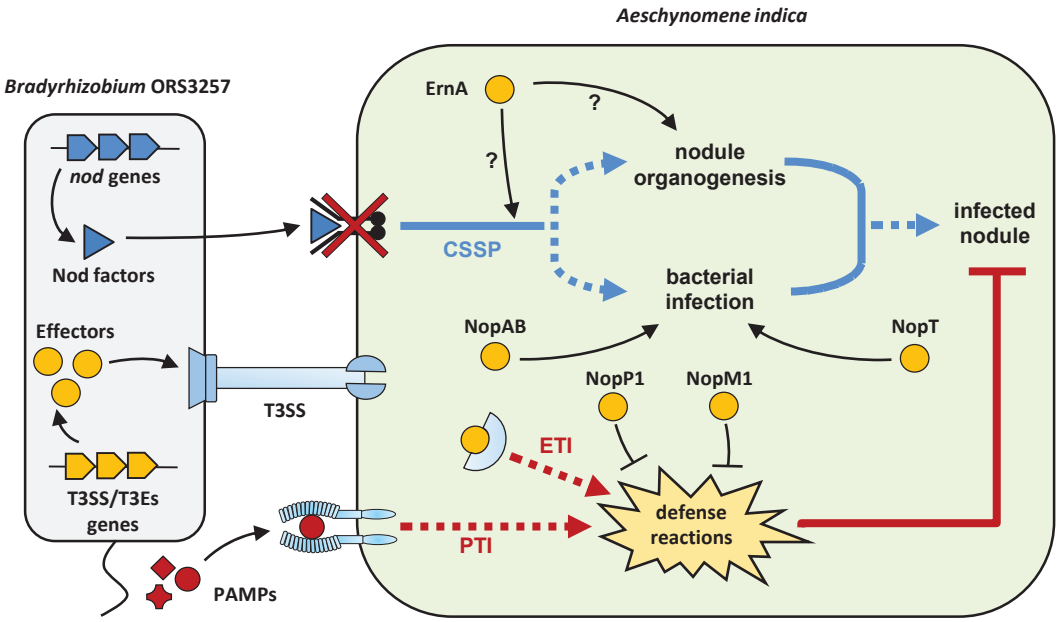


Figure 61. Modèle proposé pour la symbiose Nod-indépendante, T3SS-dépendante durant l'interaction symbiotique entre *B. vignae* ORS3257 et *A. indica* (Teulet et al., 2019).
Cette symbiose n'implique pas les facteurs Nod et dépend d'un cocktail d'effecteurs sécrétés dans la cellule hôte par le T3SS de la bactérie. Ces effecteurs agissent ensemble pour promouvoir l'organogénèse et l'infection du nodule tout en réprimant les mécanismes de défense de la plante.

teurs rhizobiens capables de déclencher directement le processus de nodulation a donc été l'une des priorités menées par les chercheurs travaillant sur cette thématique.

5.3.2. Identification de T3Es déclenchant l'organogenèse nodulaire

Pour progresser dans la compréhension du processus symbiotique Nod-indépendant, T3SS-dépendant, une mutagénèse quasi-exhaustive de l'ensemble des effecteurs prédits par des approches *in silico* chez *B. vignae* ORS3257 a été réalisée (Teulet et al., 2019). La prédiction des effecteurs a été effectuée grâce à la prédiction du motif *tts* box, retrouvé le plus souvent en amont des T3Es et sous le contrôle du régulateur transcriptionnel *ttsI* (Dong & Song, 2020) (figure 60), lui-même contrôlé par NodD, et cette prédiction a pu être confirmée par des approches protéomiques et transcriptomiques (Busset et al., 2021). Ces travaux ont permis de montrer que la symbiose entre ORS3257 et la légumineuse tropicale *A. indica* repose sur un cocktail d'au moins cinq effecteurs jouant des rôles synergiques et complémentaires (figure 61) : i) les effecteurs NopT et NopAB sont requis pour l'infection intracellulaire des nodules, ii) les effecteurs NopM1 et NopP1 jouent un rôle dans l'infection chronique des nodules et iii) un nouvel effecteur nommé ErnA[106] joue un rôle essentiel dans l'organogenèse nodulaire (Teulet et al., 2019). En effet, l'expression ectopique d'*ernA* chez des racines transformées d'*A. indica* par hairy-root[107] montre qu'en absence de la bactérie, ErnA induit la formation de méristèmes de type racinaire et nodulaire tout le long de la racine transformée. Bien qu'aucun domaine fonctionnel n'ait pu être identifié chez ErnA, il a pu être détecté la présence d'un motif NLS[108]. En accord avec cela, il a pu être montré qu'ErnA est bien adressé dans le noyau après sa translocation dans la cellule hôte (Teulet et al., 2019). De plus, par des approches de FRET-FLIM[109], il a pu être mis en évidence qu'ErnA interagit avec les acides nucléiques, mais la séquence cible ARN ou ADN reste à définir. L'ensemble de ces découvertes suggère que ce nouvel effecteur pourrait déclencher la nodulation en interférant directement sur l'expression de certains gènes de l'hôte.

Plus récemment, il a pu être montré que l'effecteur majeur gouvernant la nodulation lors de l'interaction entre la souche *B. elkanii* USDA61 et *Glycine max* cv. Enrei correspond à une SUMO-protéase annotée Bel2-5 (Ratu, Hirata, et al., 2021). Tout comme ErnA, Bel2-5 présente un motif NLS dans sa séquence et a été montré adressé au noyau. Ce nouvel effecteur partage une certaine similarité avec l'effecteur NopD identifié chez *Bradyrhizobium* sp. XS1150 mais qui a un effet négatif sur

[106] Effector Required for Nodulation A.
[107] Ou chevelu racinaire ; technique de transformation transitoire médiée par *Agrobacterium rhizogenes*.
[108] Nuclear Localisation Signal.
[109] Florescence Resonance Energy Transfer with Florescence Lifetime Imaging Microscopy. Technique qui par mesure de la durée de vie de la fluorescence de la GFP peut déterminer une association proche avec les acides nucléiques marqués au SYTOX qui est un colorant fluorescent orange.

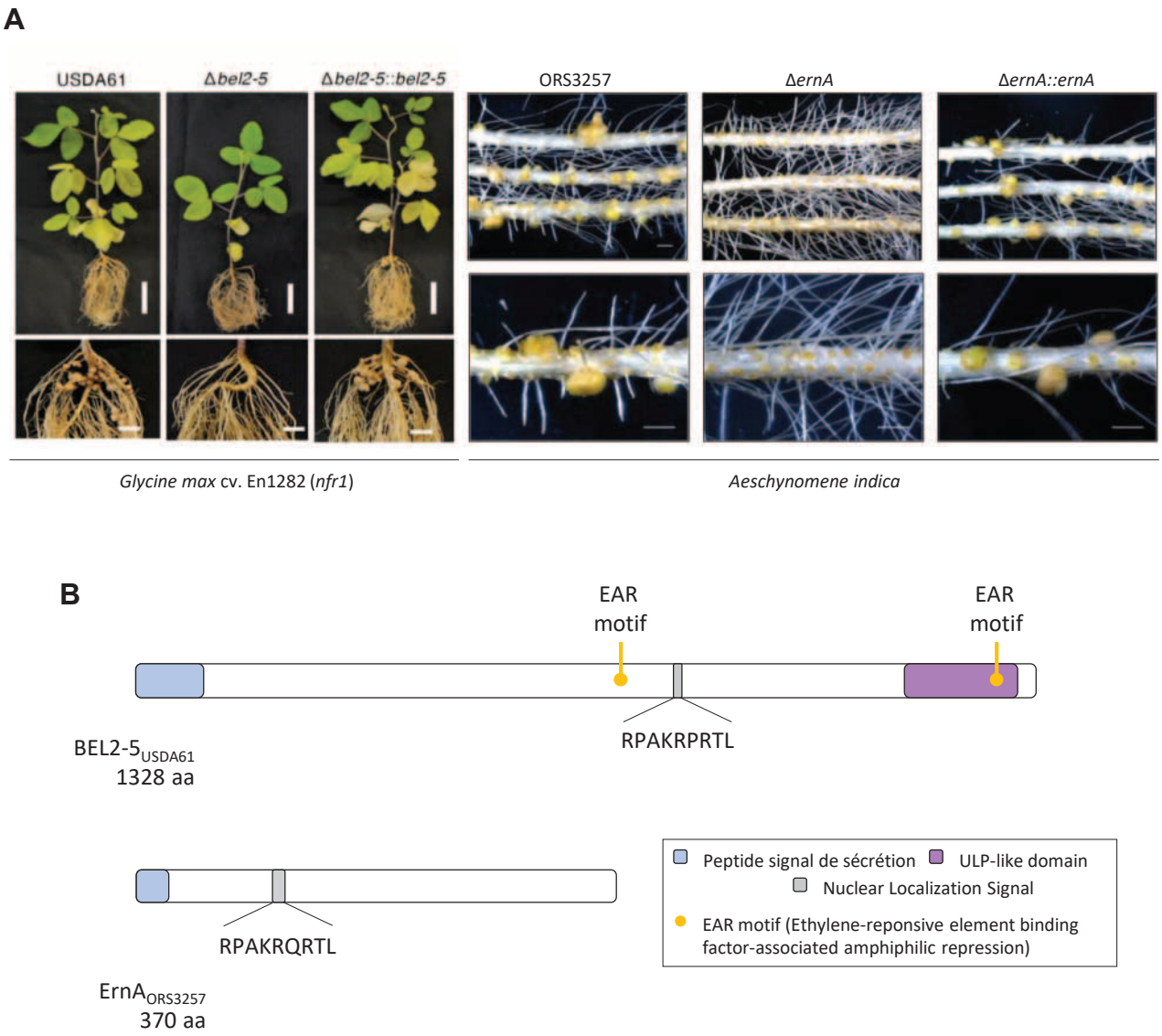


Figure 62. Bel2-5 et ErnA : deux effecteurs déclenchant la nodulation chez les légumineuses (adapté de Ratu, Teulet et al., 2021 et Teulet et al., 2019).

Ces deux effecteurs appartiennent à la famille des ET-Nod (Effector Triggering Nodulation) car étant capable d'initier l'organogénèse nodulaire. (**A**) Phénotype symbiotique des souches sauvages USDA61 et ORS3257 et de leurs mutants respectifs sur leur plante hôte (soja ou *Aeschynomene*). (**B**) Représentation schématique des effecteurs Bel2-5 et ErnA ainsi que leurs différents domaines et motifs identifiés. aa, acides-aminés.

l'efficience symbiotique avec *Tephrosia vogelii* (Xiang et al., 2020). Des similarités sont aussi retrouvées entre Bel2-5 et l'effecteur pathogène XopD identifié chez *Xanthomonas campestris* pv. *vesicatoria*[110] dont notamment par la prédiction de deux motifs EAR[111] et le domaine ULP[112] (Hotson et al., 2003; Kim et al., 2013). Il a ainsi été proposé que les rhizobiums utilisent et ont adapté des T3Es pathogènes pour promouvoir la nodulation des légumineuses (Ratu, Teulet, et al., 2021).

Par ailleurs, la présence de séquences répétées chez Bel2-5 mais dont le rôle est inconnu, semblerait être importante pour la nodulation (Ratu, Hirata, et al., 2021). Il a pu être montré que des séquences répétées en tandem telles que celles retrouvées chez les effecteurs TALE[113] (Richter et al., 2014) ou chez l'effecteur pathogène HsvG (Nissan et al., 2006) sont requises pour l'activation transcriptionnelles de gènes cibles. Ainsi il peut être supposé que ces séquences répétées retrouvées chez Bel2-5 auraient un rôle dans la reconnaissance spécifique de motifs sur l'ADN ou de protéines cibles et participerait à la régulation de gènes nécessaire à l'organogenèse nodulaire.

Ce caractère bipolaire de ce T3E peut être illustré par la capacité de Bel2-5 à déclencher la nodulation dans le contexte mutant *nfr1* (cv. En1282) alors que chez un autre cultivar (BARC-2) portant l'allèle *Rj4* il déclenche une ETI entraînant une incompatibilité d'hôte (Faruque et al., 2015; Ratu, Teulet, et al., 2021). Cette notion sera plus longuement détaillée dans la partie 6 où est présentée une revue à laquelle j'ai participée et détaillant les multiples facettes que peuvent jouer les T3Es de rhizobiums.

ET-Nod : une nouvelle classe d'effecteurs déclenchant la nodulation

Les deux effecteurs ErnA (370 aa) et Bel2-5 (1 328 aa) que nous qualifierons d'ET-Nod pour "Effectors Triggering Nodulation", ne présentent aucune homologie de séquence ce qui suggère l'existence de mécanismes moléculaires différents mis en jeu chez ORS3257 et USDA61 pour déclencher la nodulation chez *Aeschynomene* ou le soja respectivement (figure 62). La caractérisation fonctionnelle et le mode d'action de ces ET-Nods restent à définir plus précisément pour mieux comprendre les bases moléculaires de ces processus symbiotiques alternatifs reposant sur le T3SS. Cependant, l'existence de ces deux effecteurs jouant un rôle certain sur le piratage de la voie de signalisation symbiotique commune ainsi que la forte proportion des souches de *Bradyrhizobium* possédant un T3SS (Teulet et al., 2020) laisse suggérer que la symbiose Nod-indépendante, T3SS dépendante pourraient être un processus fréquemment utilisé par les *Bradyrhizobium* pour noduler diverses légumineuses. Ainsi, tout laisse à penser que la famille des ET-Nods pourrait aussi être plus large que les deux premiers représentants identifiés à ce jour.

[110] Responsable de la gale bactérienne affectant majoritairement les tomates et les poivrons.
[111] Ethylene-responsive element-binding factor-associated Amphiphilic Repression.
[112] Ubiquitin-like protease.
[113] Transcription activator-like effectors.

Effector	**Functional domains**	**Protein homology**	Subcellular localization	Function, target
ErnA	-	rhizobium-specific	Nucleus	-
GunA/GunA2	Glycoside hydrolase family 12 (GH12)	rhizobium-specific	-	Cellulase (xyloglucan hydrolase)
InnB	-	rhizobium-specific	-	-
NopAB	-	rhizobium-specific	-	Intracellular infection
NopC	-	rhizobium-specific	-	-
NopD	C48 cysteine protease (ULP)	XopD (*X. campestris*), PsvA (*P. syringae*)	Nucleus	deSUMOylation, F-Box/LRR-repeat protein
Bel2-5	C48 cysteine protease (ULP)	XopD (*X. campestris*), PsvA (*P. syringae*)	Nucleus	deSUMOylation
NopE1/NopE2	DUF1521	-	-	Reduces phytohormone
NopF	Acyl-CoA-N-acyltransferase	HopBG1 (*P. syringae*)	-	-
NopI	-	rhizobium-specific	-	-
NopJ	C55 cysteine protease, Ser/Thr acetyltransferase	YopJ/AvrRXV (*Yersinia* sp./*X. campestris*)	-	-
NopL	-	rhizobium-specific	Nucleus	MAPK substrate-SIPK, PP2C-related protein and RPK
NopM	NEL E3-ubiquitin ligase, LRR	IpaH-like (*Shigella* sp.)	Nucleus	NEL E3-ubiquitin ligase - SIPK
NopP	-	rhizobium-specific	Plasma membrane	MAPK substrate, TRAPPCC13 and TLP
NopT	C58 cysteine protease	AvrPphB/YopT (*P. syringae*/*Yesinia* sp.)	Plasma membrane	Lipid acylation modification, ATP-CSACP2 and HIRP

Tableau 3. Liste des T3Es de rhizobiums décrits à ce jour (adapté de Jimenez-Guerrero et al., 2022).
A noter que la famille des NopM (NEL-E3-ubiquitin ligase) ainsi que celle des SUMO-protéases (ULP-ubiquitin-like protease) font parties des familles les plus représentées chez les rhizobiums.

6. Revue : The versatile roles of type III secretion systems in rhizobium-legume symbioses.

Durant ma première année de thèse, j'ai eu l'opportunité de participer à l'écriture d'une revue commandée par le journal *Annual Review of Microbiology* et présentant ce qui était alors connu sur le rôle fonctionnel du T3SS durant la symbiose rhizobium-légumineuse.

Cette revue fait donc office de chapitre 6 de cette synthèse bibliographique, décrivant le T3SS depuis les gènes codant pour ses protéines de structures, jusqu'à sa régulation génétique et le rôle des T3Es sécrétés dans la cellule hôte. En complément de cette revue est ajouté un tableau des T3Es rhizobiens caractérisés à ce jour (Jiménez-Guerrero et al., 2022) (tableau 3).

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

ANNUAL REVIEWS CONNECT
www.annualreviews.org
• Download figures
• Navigate cited references
• Keyword search
• Explore related articles
• Share via email or social media

Annual Review of Microbiology

The Versatile Roles of Type III Secretion Systems in Rhizobium-Legume Symbioses

Albin Teulet,[1,*] Alicia Camuel,[1,2,*] Xavier Perret,[3,*] and Eric Giraud[1,2,*]

[1]Laboratoire des Symbioses Tropicales et Méditerranéennes (LSTM), IRD, Institut Agro, INRAE, Université de Montpellier, and CIRAD, Montpellier, France; email: eric.giraud@ird.fr

[2]PHIM Plant Health Institute, IRD, Institut Agro, INRAE, Université de Montpellier, and CIRAD, Montpellier, France

[3]Laboratory of Microbial Genetics, Department of Botany and Plant Biology, University of Geneva, Geneva, Switzerland

Annu. Rev. Microbiol. 2022. 76:45–65

First published as a Review in Advance on April 8, 2022

The *Annual Review of Microbiology* is online at micro.annualreviews.org

https://doi.org/10.1146/annurev-micro-041020-032624

Copyright © 2022 by Annual Reviews. All rights reserved

*These authors contributed equally to this article

Keywords

T3SS, effector, symbiosis, rhizobium, legume, plant immunity, nodule

Abstract

To suppress plant immunity and promote the intracellular infection required for fixing nitrogen for the benefit of their legume hosts, many rhizobia use type III secretion systems (T3SSs) that deliver effector proteins (T3Es) inside host cells. As reported for interactions between pathogens and host plants, the immune system of legume hosts and the cocktail of T3Es secreted by rhizobia determine the symbiotic outcome. If they remain undetected, T3Es may reduce plant immunity and thus promote infection of legumes by rhizobia. If one or more of the secreted T3Es are recognized by the cognate plant receptors, defense responses are triggered and rhizobial infection may abort. However, some rhizobial T3Es can also circumvent the need for nodulation (Nod) factors to trigger nodule formation. Here we review the multifaceted roles played by rhizobial T3Es during symbiotic interactions with legumes.

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

Contents

INTRODUCTION 46
CHARACTERISTICS SPECIFIC TO RHIZOBIAL T3SSs 47
IN RHIZOBIA, T3SSs AND NODULATION GENES ARE COREGULATED ... 48
DISTRIBUTION AND EVOLUTION OF T3SSs IN RHIZOBIA 50
NO UNIVERSAL RECIPE FOR COCKTAILS OF RHIZOBIAL EFFECTORS... 51
SYMBIOTIC ROLES FOR RHIZOBIAL T3Es 51
FUNCTIONAL CHARACTERIZATION OF T3Es 53
Targeting the Ubiquitin-SUMO Pathways of Hosts 53
Modulation of MAPK Signaling 54
Proteolytic Activities of Rhizobial T3Es 54
TARGETS OF RHIZOBIAL T3Es 54
T3Es AND EFFECTOR-TRIGGERED IMMUNITY IN LEGUMES 55
T3SS ALTERNATIVES TO NOD FACTOR–DEPENDENT NODULATION ... 56
RHIZOBIAL T3Es CAPABLE OF TRIGGERING NODULATION 57
CONCLUSIONS 57

INTRODUCTION

Rhizobia form a paraphyletic group of alpha- and betaproteobacteria able to activate formation of nitrogen-fixing nodules on the roots of legumes. Once sheltered inside nodule cells, rhizobia provide reduced forms of atmospheric nitrogen (N_2) in exchange for various nutrients (92). Ammonium provided by rhizobia allows legumes to grow in nitrogen-poor soils, thus limiting the need for chemical fertilizers to boost crop yields. Associations between rhizobia and legumes are extremely diverse and occur naturally in many terrestrial ecosystems; as such, they are considered to be major contributors of fixed nitrogen and hence of prime importance for sustainable and environmentally friendly agriculture (30).

Soils are complex ecosystems in which plant roots interact with a broad diversity of microbes, some of which are pathogens or symbionts (73, 89). How legumes secure the entry of symbiotic rhizobia while preventing infections by pathogens remains one of the most burning questions about plant-microbe interactions (42). In general, the symbiotic process begins when rhizobia secrete strain-specific nodulation (Nod) factors (NFs) in response to flavonoids exuded by roots. Perception of NFs by plant receptors activates cellular programs leading to nodule organogenesis and the formation of infection threads that guide infecting bacteria to the emerging nodule primordium (86, 88). When a symbiosis is proficient, each of the infected nodule cells may harbor thousands of nitrogen-fixing bacteroids. How rhizobia can establish such large intracellular communities is intriguing given that plants have evolved potent immune responses to repel microbial infections (6). In plants, recognition of microorganisms is mediated by membrane receptors that detect microbe- or pathogen-associated molecular patterns (MAMPs or PAMPs) and that activate a first line of defense called MAMP- or PAMP-triggered immunity (MTI/PTI) (10, 134).

Although no single rhizobium-specific MAMP has been identified as acting on legumes so far, transient defense-like reactions consistent with MTI responses are often observed during the early stages of infection (39, 56, 63, 128). While NFs have been shown to attenuate MTI (28, 61), rhizobia use additional signals and mechanisms to facilitate infection of legume roots. One of the most striking rhizobial strategies to promote legume infection is the use of effector proteins delivered by type 3 secretion systems (T3SSs), which subvert plant immunity in ways similar to mechanisms

of pathogenic bacteria (21). Initially identified in animal- and plant-pathogenic bacteria, T3SSs deliver cocktails of effector proteins (T3Es) inside the cytoplasm of host cells. T3Es interfere with various cellular processes to sabotage immune responses and promote bacterial invasions (15, 24, 31). To defend against these T3SS-mediated attacks, plants have evolved a family of cytoplasmic receptors that, once associated with cognate T3Es, trigger a powerful second line of defense called effector-triggered immunity (ETI) (18). ETI is an extremely potent, hypersensitive response often associated with cell death that halts microbial infection and helps render plants resistant to pathogens (51).

Despite the potential benefits that rhizobia may represent for legumes, a growing body of literature indicates that T3Es can also compromise symbiosis through ETI, as already reported for plant-pathogen interactions (74, 142, 143). This makes rhizobial T3SSs double-edged swords that may promote or prevent symbiosis depending on the host plant (114). Interestingly, recent studies revealed that rhizobial T3SSs can also trigger nodule organogenesis directly, thus bypassing the need for compatible NFs (84, 85, 93, 121). Here, we review the many facets of T3SSs in the context of the beneficial symbioses between rhizobia and legumes.

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

CHARACTERISTICS SPECIFIC TO RHIZOBIAL T3SSs

T3SS secretory machinery, also called the injectisome, is composed of about 20 different proteins that form a basal body across the bacterial envelope, and an extracellular needle that connects the bacterium to the eukaryotic cell targeted for T3E secretion (**Figure 1**). While the structure of T3SSs of different human pathogens such as *Salmonella*, *Yersinia*, and *Escherichia* species has been extensively studied (14, 24, 83), the corresponding secretion apparatuses of rhizobia remain poorly characterized. Given the high degree of conservation between proteins that form the basal bodies of rhizobia and pathogen T3SSs, the secretion machineries are expected to share similar ultrastructures and functions. However, rhizobial T3SSs also have specific features such as a secretin, which is split in the RhcC1 and RhcC2 components (1).

Like those observed for plant pathogens, the T3SS pili of rhizobia are thin and up to 2-μm-long, flexible extracellular microfilaments (22, 59, 66, 98). In contrast to the 40- to 80-nm needles of T3SSs of animal pathogens, the longer T3SS appendages of rhizobia are possibly adapted to cross plant cell walls and/or result from the absence of molecular rulers, such as the *Yersinia* YscP and *Pseudomonas* PscP proteins, which were shown to control the length of needles (4, 14).

Isolation and partial purification of T3SS appendages showed that pili from *Ensifer* (formerly *Sinorhizobium*) species are made of at least three nodulation outer proteins (Nops) that cross-interact and are specific to rhizobia: NopA, NopB, and NopX (22, 58, 66, 98, 99). It has been suggested that NopA and NopB pilins are part of the extracellular filament, whereas NopX proteins may form translocon structures across the host cell membrane (99). Interestingly, several bradyrhizobia with functional T3SSs, such as the model strain *Bradyrhizobium diazoefficiens* USDA110, were shown to lack a *nopX* homolog (57, 118). Since a recent large-scale comparative analysis of bradyrhizobia genomes showed that strains that lacked *nopX* instead carried T3SS-related *nopE* and *nopH* genes, it is tempting to speculate that products of the two latter genes functionally replace NopX in the translocation of T3Es (122).

As the extracellular appendage of rhizobial T3SSs may come in direct contact with host cells, pilus proteins may act as MAMPs. Indeed, the HrpE pilin of the plant pathogen *Xanthomonas campestris* was shown to elicit MTI (38, 105). The different protein compositions of T3SS pili may therefore correspond to adaptations needed for bacteria to evade plant immune responses or, conversely, to favor recognition by specific hosts (14). Whether T3SS pilins of rhizobia evolved to avoid host immune responses, as already proposed for their flagellins (37, 39, 65), remains to be tested.

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

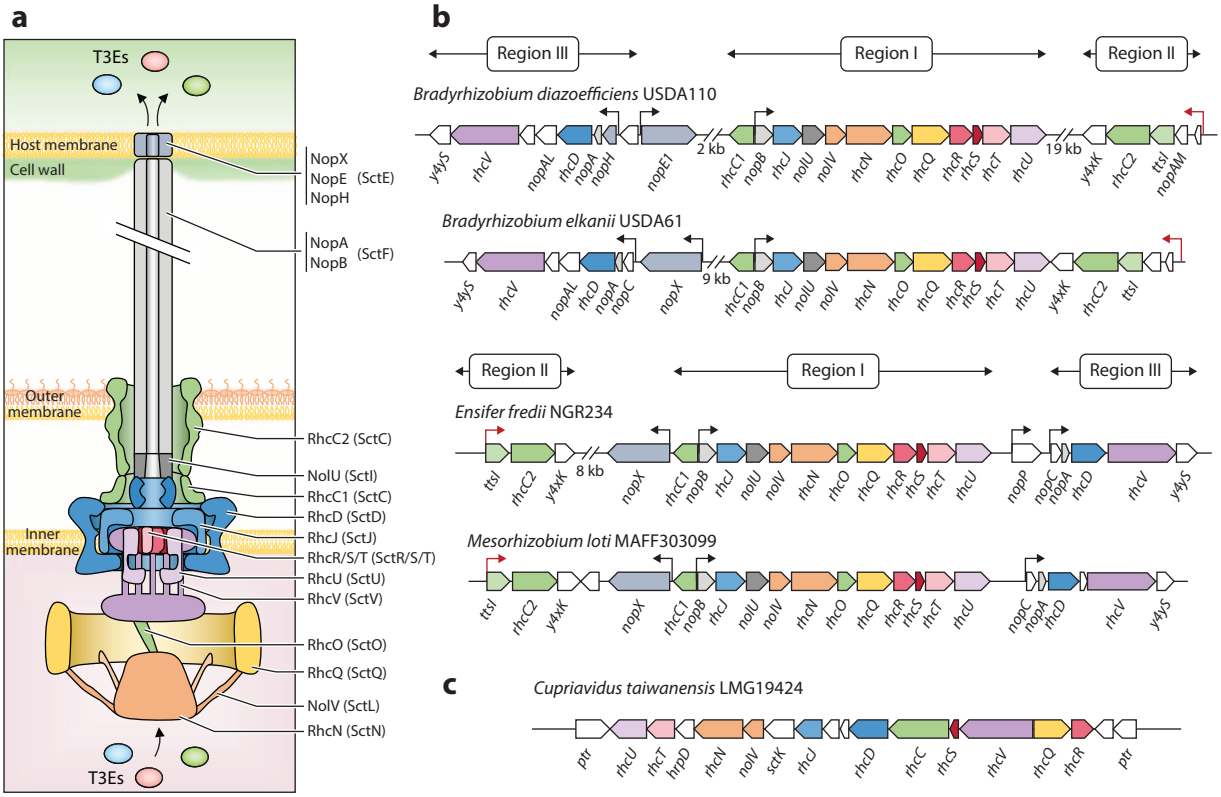


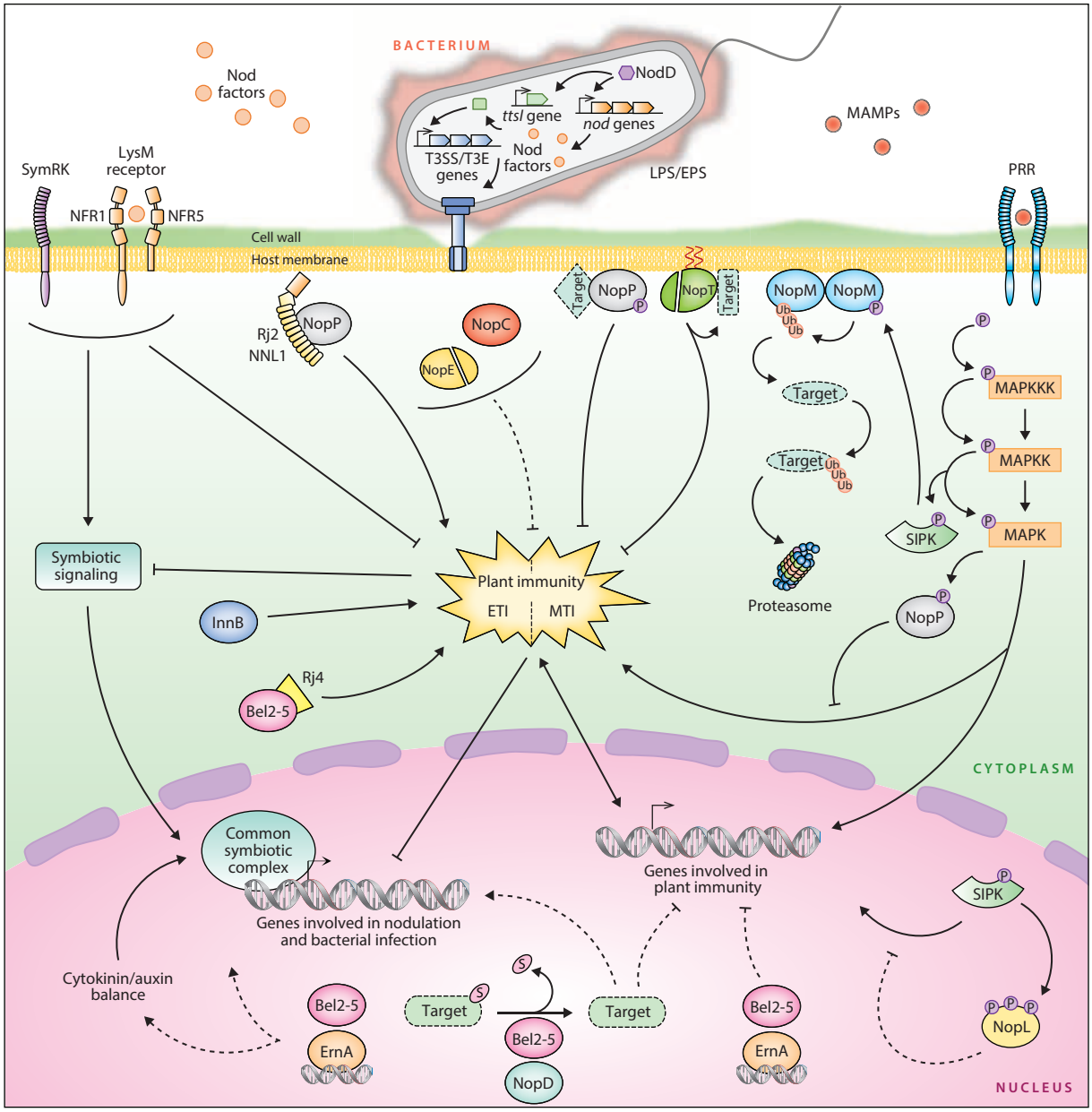
Figure 1

Schematic representation of the type III secretion system (T3SS) machinery and genetic organization of symbiotically active T3SS gene clusters identified in rhizobia. (*a*) Proteins that form the T3SS apparatus are named according to the established rhizobia nomenclature. The corresponding unified Sct (secretion and cellular translocation) names are in parentheses. Panel *a* adapted from References 114 and 130. (*b*) Genetic organization of T3SS clusters identified in various α-rhizobia with three mostly conserved regions (regions I to III) as defined by Tampakaki (118). Genes are represented by arrows outlined in black, oriented according to transcription, and colored according to the colors of the corresponding proteins in panel *a*. White arrows represent genes whose products are type III effectors (T3Es) (e.g., NopP) or of unknown function (e.g., y4yS). Thin black and red arrows indicate promoters with conserved *tts* and *nod* boxes, respectively. (*c*) Atypical T3SS gene cluster identified in the β-rhizobium *Cupriavidus taiwanensis* LMG19424, with its putative transcriptional regulator (*ptr*).

IN RHIZOBIA, T3SSs AND NODULATION GENES ARE COREGULATED

Genes coding for structural components of T3SSs are clustered in rhizobia and are often found close to symbiotic genes like the *nod* genes needed for synthesis of NFs. In α-rhizobia, except for the *rhcC1* secretin and the putative translocators *nopX* and *nopE* monocistronic genes, most T3SS genes are organized in three major genetic units thought to form distinct operons (118, 122) (**Figure 1**). One major operon containing most of the T3SS structural genes is highly conserved, whereas the two remaining operons often display strain-specific differences in gene composition and genetic organization. In β-rhizobia, such as *Cupriavidus taiwanensis* strain LMG19424, genes encoding components of a symbiotically active T3SS are organized in two divergently transcribed units that are more similar in genetic organization to the human opportunistic pathogen *Burkholderia cenocepacia* than they are to the T3SS loci of α-rhizobia (97) (**Figure 1**).

A remarkable feature of α-rhizobial T3SS genes is their coregulation with major *nod* genes, whose transcription is flavonoid- and NodD-dependent. In response to plant-made flavonoids and via binding to specific *nod* box regulatory sequences, transcription regulators of the NodD family activate *nod* genes as well as *ttsI* (57, 129) (**Figure 2**). Once expressed, the transcriptional activator TtsI binds to conserved promoter sequences called *tts* boxes found upstream of genes and operons coding for T3SS components and T3Es (11, 57, 64, 70, 135, 144). Interestingly, TtsI may regulate functions other than those needed for T3SSs, like in *Ensifer fredii* strain NGR234,



(*Caption appears on following page*)

Erratum >

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

Figure 2 (*Figure appears on preceding page*)

The multifaceted roles of rhizobial T3SSs in legume symbioses. Molecular interplay between factors secreted by rhizobia and some of the plant components involved in perceiving microbes. Plants can detect bacteria via membrane receptors (PPRs) that recognize specific MAMPs and activate a first line of defense (MTI). Concomitantly, in compatible rhizobia, plant flavonoids trigger the synthesis and secretion of NFs and T3Es, via the successive activation of NodD and TtsI transcription regulators. Membrane-bound LysM-RLK receptors (NFR1 and NFR5) associated with SymRK and compatible NFs activate the symbiotic signaling pathway, leading to nodule organogenesis and bacterial infection. T3Es translocated into the host cell cytoplasm by rhizobial T3SSs modulate the symbiotic outcome depending on the host plant. Some T3Es, including NopM, NopP, NopD, NopL, and NopT, interfere with different components of the plant immune systems and thus promote symbiosis by repressing MTI responses. Alternatively, when detected by cytosolic receptors (Rj4, Rj2, GmNNL1), some T3Es (NopP, InnB, Bel2-5) trigger ETI-type responses that block nodulation. Eventually, T3Es such as ErnA and Bel2-5 may directly trigger nodule organogenesis via an unknown mechanism possibly involving binding to nuclear DNA. Note that the figure summarizes a set of different responses mediated by rhizobial T3Es in different plant species. Arrows represent documented interactions. Dashed lines represent hypothetical functions. Abbreviations: EPS, exopolysaccharide; ETI, effector-triggered immunity; LPS, lipopolysaccharide; LysM-RLK, lysin motif receptor-like kinase; MAMP, microbe-associated molecular pattern; MAPK, mitogen-activated protein kinase; MAPKK, MAPK kinase; MTI, MAMP-triggered immunity; NF, Nod factor; P, phosphorylated site; PRR, pattern recognition receptor; S, SUMOylation; SIPK, salicylic acid–induced protein kinase; T3E, type III effector; T3SS, type III secretion system; Ub, ubiquitination.

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

where TtsI also activates loci needed for modifying lipopolysaccharides (LPS) into a symbiotically active form called rhamnan (70, 95). In contrast, activation of T3SS functions in *C. taiwanensis* was reported to be independent of both flavonoids and TtsI, whereas it was triggered by glutamate, like in plant pathogens such as *Ralstonia solanacearum* and *Pseudomonas aeruginosa* (97).

DISTRIBUTION AND EVOLUTION OF T3SSs IN RHIZOBIA

Rhizobia form a large polyphyletic group of hundreds of species distributed in as many as 18 genera of alpha- and betaproteobacteria (20; **https://www.bacterio.net**). The propensity of nodulation (*nod*) and nitrogen fixation (*nif* and *fix*) genes to be carried by genomic islands and plasmids that can be transferred laterally between strains is thought to greatly contribute to the existing diversity of rhizobia (72, 92, 94). Symbiotic tool kits do not necessarily include T3SSs, whose prevalence varies significantly among genera. In the major *Ensifer*, *Mesorhizobium*, and *Rhizobium* genera, co-occurrence of *nod* and T3SS genes is not the rule (115, 125, 133). For example, among 48 *Ensifer* strains analyzed, only a third possessed a T3SS, but all were found to carry type IV secretion systems (T4SSs) (115). Like T3SSs, T4SSs deliver effector molecules to the cytoplasm of host cells (17, 40), and several reports have confirmed that T4SSs of rhizobia also modulate symbiosis positively or negatively, depending on the legume host (21, 45, 77). By contrast, more than 90% of the nodule isolates that belong to the *Bradyrhizobium* genus carry conserved *nod* and T3SS genes that are embedded in symbiotic islands; phylogenetic analyses indicated that these symbiotic genes share a common evolutionary history (122). As several reports have proposed bradyrhizobia to be the ancestors of all rhizobia (46, 87, 133), the transformation of a free-living bacterium into an endosymbiont of legumes may have required both NFs and T3Es to achieve, respectively, nodule formation and subversion of host immunity to secure intracellular infection of nodule cells. If so, during their coevolution with their hosts, rhizobia may also have developed alternatives to T3SSs to cope with plant immunity, for example, by modifying cell surface components (e.g., exopolysaccharides, LPS, or K-antigen polysaccharides) or by using alternative protein secretion systems (34, 39, 47, 78, 92).

Of the seven families of T3SS genetic clusters in gram-negative bacteria that have been cataloged (32, 126), rhizobia primarily use the α-RhcI type to target legume hosts. Some strains possess up to two apparently complete additional atypical T3SS clusters; however, little is known about their function(s) (118, 122). In the *E. fredii* strain NGR234, inactivation of the T3SS-II locus had no measurable effect on symbiosis (103). By contrast, although inactivation of the atypical T3SS of

C. taiwanensis strain LMG19424 had no effect on its primary host, *Mimosa pudica*, it extended the host range of the mutant to include *Leucaena leucocephala* (97). It is possible that these additional atypical T3SSs provide rhizobia with a selective advantage, for example, during colonization of the host rhizosphere and/or competition against other soil microbes or organisms.

NO UNIVERSAL RECIPE FOR COCKTAILS OF RHIZOBIAL EFFECTORS

Approaches based on secretome analysis enabled the first identifications of rhizobial T3Es in *Ensifer* and *Bradyrhizobium* strains (43, 71, 129). Proteins secreted by flavonoid-activated T3SSs were originally called nodulation outer proteins (Nops), to mirror the nomenclature of the *Yersinia* outer proteins (Yops) (69). However, not all Nops should be considered bona fide effectors since NopA, NopB, and NopX are pilus components, and conversely, some of the recently identified T3Es, such as InnB, Bel2-5, and ErnA, were not designated as Nops.

Actually, few of the secreted proteins have been confirmed to possess all the hallmarks of genuine symbiotic T3Es: (*a*) secretion by a T3SS, (*b*) translocation into host cells, and (*c*) modulation of symbiotic interactions with legume hosts. Thus far, only NopC, NopL, and NopP of *E. fredii* strains; NopE1 and NopE2 of *B. diazoefficiens* USDA110; Bel2-5 and InnB of *Bradyrhizobium elkanii* USDA61; and ErnA of *Bradyrhizobium vignae* ORS3257 meet all three criteria for bona fide T3Es (13, 49, 50, 79, 93, 100, 121).

Whereas homology searches for T3SS machinery proteins are simple, identifying the complete set of T3Es (effectome) of a symbiotic strain is more difficult. Genome mining analyses based on combined searches for regulatory *tts* boxes and/or homologies to known T3Es of symbiotic or pathogenic bacteria helped define the putative effectomes of rhizobia (13, 55, 91, 93, 109, 110, 121, 122). Although these *in-silico* searches may occasionally overestimate the number of T3Es secreted by rhizobia (13), comparison of the predicted effectomes revealed interesting properties. The size of T3E repertoires varies considerably between strains, from circa 10 in *Ensifer* strains to more than 50 for some bradyrhizobia such as *B. elkanii* USDA61 (55, 93). Although no core effectome has been identified in rhizobia, NopC, NopM, NopP, and NopT homologs are frequently shared between *Bradyrhizobium* and *Ensifer* strains, suggesting the existence of a possible early effectome core (122). Several of the predicted T3Es, such as ErnA, NopC, NopL, NopP, and NopI, appear to be specific to rhizobia (21, 49, 50, 121). Such differences in the composition of the predicted effectomes indicate high plasticity and rapid evolution of the T3E arsenals in rhizobia, possibly for a better adaptation to specific hosts.

SYMBIOTIC ROLES FOR RHIZOBIAL T3Es

Several studies have highlighted that the deletion of specific T3Es may improve, impair, or have no effect on the symbiotic outcome, depending on host plants (**Figure 3**). For example, compared to the parent *E. fredii* strain NGR234, a *nopT* mutant formed fewer nodules on *Tephrosia vogelii* but nodulated *Crotalaria juncea* better (19, 52), while remaining as proficient as the wild type on *Lablab purpureus* (52). Although listing all of the symbiotic properties of T3E mutants is beyond the scope of this review, it is noteworthy that host-dependent positive or negative effects on symbiosis were also reported when each of the following putative effector genes was mutated: *bel2-5*, *ernA*, *nopAB*, *nopC*, *nopD*, *nopE*, *nopF*, *nopI*, *nopJ*, *nopL*, *nopM*, *nopP* and *innB*, (33, 48–50, 52, 60, 79, 90, 93, 108, 122, 136, 139). A cumulative effect on symbiosis of deleting multiple effectors has been reported in several strains, including in *E. fredii* NGR234, where deletion of both *nopL* and *nopP* reduced *Flemingia congesta* nodulation more than mutation of either of these genes (107). Similarly, in *B. vignae* ORS3257, mutagenesis of both *nopP1* and *nopM1* impacted nodulation of *Aeschynomene indica* and *Vigna mungo* significantly more than single mutations of *nopP1* or *nopM1* (108, 121). It

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

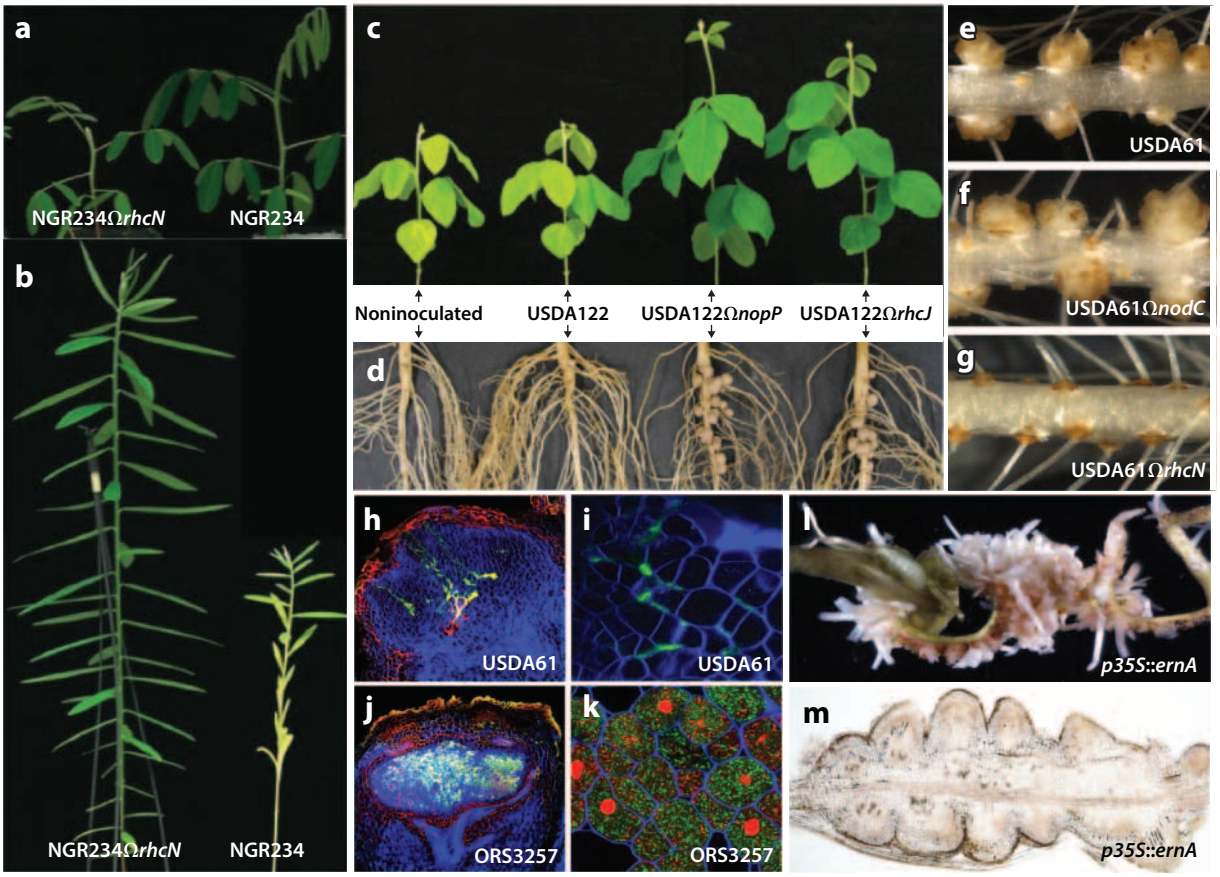


Figure 3

Impacts of T3SSs on rhizobium-legume symbioses. (*a*) In some hosts, a T3SS and its secreted T3Es can promote symbiosis, like that observed during interactions between *Ensifer fredii* NGR234 and *Tephrosia vogelii*, with plants showing reduced growth when inoculated with T3SS-deficient strains (e.g., NGR234Ω*rhcN*). (*b*) Alternatively, T3SSs may completely block symbiotic interactions, as illustrated here in the parent strain *E. fredii* NGR234 in *Crotalaria juncea*. The T3SS-mutant NGR234Ω*rhcN* is fully proficient. (*c*,*d*) One specific effector may suffice to trigger ETI-based incompatibility, as is the case of Rj2-soybean cultivars capable of detecting NopP of *Bradyrhizobium diazoefficiens* USDA122. (*e*–*g*) The T3Es of some rhizobia may also bypass the requirement for NFs to trigger nodulation, as shown in a *nodC* mutant of strain *Bradyrhizobium elkanii* USDA61 (USDA61Ω*nodC*) that is unable to secrete NFs but nonetheless forms nodules on *Aeschynomene indica* and in a similar way to the parent USDA61, while a T3SS mutant (USDA61Ω*rhcN*) fails to nodulate. Depending on the strain, the nodules elicited on *A. indica* via the T3SS are infected (*h*,*i*) intercellularly or (*j*,*k*) intracellularly. (*l*,*m*) The ErnA effector of *Bradyrhizobium vignae* ORS3257 confers the ability to form nodules, as shown by ectopic expression of ErnA in transgenic lines of *A. indica*, which results in formation of nodule-like structures. Abbreviations: ETI, effector-triggered immunity; NF, Nod factor; T3E, type III effector; T3SS, type III secretion system. Panels *a* and *b* adapted from Reference 107; panels *c* and *d* adapted from Reference 116 (CC BY 4.0); panels *e*–*k* adapted from Reference 85; and panels *l* and *m* adapted from Reference 121.

is thus now firmly established that in general, the role of rhizobial T3SSs in symbiosis does not depend on a single effector but rather results from the combined activities of a cocktail of secreted T3Es that can act synergistically, redundantly, and/or antagonistically, depending on the legume host that is challenged (21, 108).

In animal pathogens, effectors are known to be translocated through T3SSs sequentially with chaperone proteins controlling both the order and the time of secretion (14, 127, 138). A study of

the T3SS-mediated interactions between *B. elkanii* USDA61 and several *Lotus* accessions showed that NopF inhibited the infection of *L. japonicus* cultivar Gifu, whereas NopM was responsible for inducing early senescence in nodules of *L. japonicus* cultivar MG-20 (60). Although sequential secretion of T3Es in USDA61 cannot be excluded, no T3SS-dependent chaperone has been identified in rhizobia. Since in vitro, genistein-induced cells of USDA61 secreted NopF and NopM together with other T3Es (60), an alternative hypothesis to sequential secretion of T3Es is that plant cells that are either crossed by infection threads or located inside the nodules differ in susceptibility to T3Es.

FUNCTIONAL CHARACTERIZATION OF T3Es

Like T3Es of pathogenic bacteria, rhizobial T3Es often display a modular architecture with different functional domains and motifs for subcellular localization once T3Es are translocated into host cells (3, 25, 114). Only a few rhizobial effectors have been functionally characterized, mainly using transgenic or transfected nonlegume hosts such as *Nicotiana benthamiana*, *Nicotiana tabacum*, or *Arabidopsis thaliana*, three of the most widely used models for analyzing T3Es. These experimental approaches highlighted functional homologies between T3Es of rhizobia and plant pathogens, in particular for domains needed to subvert the posttranslational modification pathways of the hosts, for example by SUMOylation, ubiquitinylation, or phosphorylation (**Figure 2**).

Targeting the Ubiquitin-SUMO Pathways of Hosts

Ubiquitin and small ubiquitin-like modifier (SUMO) are structurally related small proteins used by eukaryotic cells to specifically modify various substrate proteins. These posttranslational modifications (ubiquitinylation and SUMOylation) can influence several aspects of the protein biology, including stability, activity, location, and interactions with different partners (5, 75, 104). Although SUMO and ubiquitin are both conjugated to target proteins via three-step enzymatic cascades, distinct enzymes are involved in each step: E1 activating enzymes, E2 conjugating enzymes, and E3 ligases. Effectors of the NopM family, found in many rhizobia, carry a C-terminal novel E3 ubiquitin ligase (NEL) domain as well as a conserved N-terminal leucine-rich repeat (LRR) domain that is thought to define substrate specificity (3, 96). When expressed in *N. benthamiana*, NopM of strain NGR234 represses the production of reactive oxygen species induced by flagellin (140). Subsequent in vitro assays showed that NopM forms unanchored polyubiquitin chains and possesses auto-ubiquitination activities, thereby confirming the functionality of its NEL domain (141). Considering that most ubiquitinated proteins are degraded by 26S proteasomes (104), NopM effectors may use the ubiquitin-proteasome pathway to degrade plant proteins possibly involved in defense responses.

Unlike the NopM E3 ubiquitin ligase, other rhizobia effectors are predicted to carry a C-terminal ubiquitin-like protease (ULP)-like domain for deSUMOylation of proteins (93, 139). ULP-like domains are also conserved in several effectors of plant pathogens, including XopD of *Xanthomonas campestris* and PsvA of *Pseudomonas syringae* (54). Interestingly, NopD of *Bradyrhizobium* sp. XS1150 and Bel2-5 of *B. elkanii* USDA61, which both carry a ULP-like domain, share similarity with XopD. NopD and Bel2-5 were both shown to target the plant cell nucleus and, in their corresponding ULP-like domain, to carry the catalytic triad residues H/G/C that are required for activity (44, 53, 54, 93, 139). As NopD of strain XS1150 processed plant SUMO proteins and cleaved SUMO-conjugated substrates, the functionality of its ULP-like domain was confirmed, strengthening the view that NopD homologs act in similar ways to their plant pathogen counterparts by deSUMOylating the targeted plant proteins (139).

Notably, a search for ULP-like and NEL domains in the deduced proteomes of fully sequenced members of the *Bradyrhizobium* genus indicated that NopM and SUMO proteases are abundant in

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

symbiotic strains with a T3SS and form the two largest families of T3Es (122). Thus, the ubiquitin and SUMO modification pathways could be the plant cellular functions preferentially targeted by T3Es of bradyrhizobia during symbiosis.

Modulation of MAPK Signaling

Perception of MAMPs by plant plasma membrane receptors triggers plant immune responses via activation of mitogen-activated protein kinases (MAPKs) (134). By mimicking substrates for activity of MAPKs, many T3Es of pathogens interfere with the phosphorelays needed for PTI activation (15, 23, 134). In *E. fredii* NGR234, NopL, NopM, and NopP were shown to be substrates for plant kinases in vitro (2, 33, 107, 141), with NopM and NopL also shown to be phosphorylated *in planta* by the MAPK salicylic acid–induced protein kinase (SIPK) of *Nicotiana tabacum* (NtSIPK) (33, 140, 141, 147). Physical interactions of NopL and NtSIPK in the plant cell nucleus were also reported (33). Disruption of signaling of MAPKs in legume cells is probably the main function of NopL, since inactivation of all its multiple phosphorylation sites resulted in a symbiotic phenotype similar to that of a *nopL* knockout mutant, with premature nodule senescence in *Phaseolus vulgaris* (33). In contrast to NopL, which can be hyperphosphorylated (33), NopM carries a single serine residue found to be phosphorylated in vitro as well as in vivo. As NopM is also an E3 ubiquitin ligase, it may interfere with both MAPK signaling and protein degradation processes.

Proteolytic Activities of Rhizobial T3Es

Several effectors of plant pathogens have been shown to cleave proteins once inside host cells (15, 23). Among them, the C58 cysteine protease AvrPphB of *P. syringae* was shown to undergo self-cleavage and also to cleave several plant receptor-like cytoplasmic kinases, such as BIK1 and PBS1 of *A. thaliana*, which are involved in transducing PTI and ETI responses, respectively (25, 146). By homology, NopT effectors of rhizobia belong to the large AvrPphB/YopT family of T3Es. Several studies have indicated that, once translocated into a host cell, NopT undergoes proteolytic self-cleavage that exposes residues required for lipid acylation modification (N-myristoylation and S-palmitoylation), which are needed to subsequently target the plant plasma membrane (25, 29, 67). The typical catalytic triad residues C/H/D of cysteine proteases that were identified in NopT effectors were shown to be required for optimal self-cleavage as well as for induction of a hypersensitive response in transformed tobacco cells (19, 25, 29, 52, 67). Interestingly, the sites for autoprocessing in AvrPphB and for the AvrPphB-dependent cleavage of PBS1 share identical amino acids (25). By analogy, the DKM amino acid motif that is needed for self-cleavage of NopT of strain NGR234 is possibly also conserved in the plant proteins targeted by NopT.

When incubated in vitro in the presence of Ca^{2+}, the purified NopE1 proteins of *B. diazoefficiens* USDA110 also displayed autoproteolytic activity, but via two metal ion-inducible autocleavage (MIIA) domains (102, 136). As a noncleavable form of NopE1 did not complement a *nopE1 nopE2* double mutant of USDA110 to restore the wild-type detrimental symbiotic phenotype on *Vigna radiata* plants (136), the self-cleavage property of NopE1 appears to be required for its activity *in planta*. Whether NopE1 and NopE2 of USDA110 act as bona fide effectors by cleaving specific target proteins or play a role in the translocation of rhizobial T3Es across the host plasma membrane as proposed by Teulet et al. (122) remains to be determined.

TARGETS OF RHIZOBIAL T3Es

Besides MAPKs such as SIPK that are targeted by NopL and NopM, few other legume proteins have been identified as specific substrates for rhizobial T3E activity. Recently, *Robinia pseudoacacia*

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

proteins susceptible to being targeted by NopT and NopP of *Mesorhizobium amorphae* CCNWGS0123 were identified (62, 67). In yeast two-hybrid combinations of bait (NopT or NopP) and prey (cDNA of *R. pseudoacacia*), the membrane-associated ATP-CSACP2 (ATP-citrate synthase alpha chain protein 2) and HIRP (hypersensitive-induced response protein) interacted with NopT, while a TRAPPC13 (trafficking protein particle complex subunit 13)-like protein interacted with NopP. Although interaction of NopT and NopP with these plant proteins has been validated by bimolecular fluorescence complementation assays in *N. benthamiana*, further studies are needed to determine how these T3E/target(s) interactions modulate the host cellular processes and symbiotic properties of secreting strains.

Quantitative trait locus (QTL) analyses of modifications in the number and dry weight of nodules on soybean plants also led to the identification of possible targets for NopD, NopL, and NopP of *E. fredii* HH103 (131, 132, 148). QTL mapping in soybean showed that NopD was associated with a locus coding for one F-Box/LRR-repeat protein (132), NopL was associated with two soybean genes coding for a protein phosphatase 2C (PP2C)-related protein and a receptor protein-tyrosine kinase (RPK) (148), and NopP was associated with a thaumatin-like protein (TLP) and MAPK3 (131). While the putative functions assigned to these candidate loci are consistent with T3Es interfering with plant immunity, further studies are needed to confirm whether the interactions between NopD, NopL, and NopP and the candidate plant proteins are direct or indirect.

T3Es AND EFFECTOR-TRIGGERED IMMUNITY IN LEGUMES

An increasing number of reports confirm that rhizobial T3Es can trigger ETI responses in legumes, thus restricting the host range of strains and/or limiting symbiotic proficiency to specific cultivars (74, 114). Products of dominant alleles have been known for decades to restrict nodulation of specific strains of *Glycine max* (16). Yet, genetic analyses only recently revealed the roles in host specificity of cytosolic plant receptors and of the corresponding T3Es. For example, the *Rj2* and *Rfg1* alleles of the same locus on soybean chromosome 16 code for nearly identical 1,052-residue-long Toll-interleukin receptor (TIR)-nucleotide-binding site (NBS)-LRR proteins involved in the contrasted specificities of *Bradyrhizobium japonicum* USDA122 and *E. fredii* USDA257 (142). Inoculation of the *Rj2*-genotype *G. max* cultivar Hardee with natural variants of strain USDA122 showed that NopP determines the symbiotic incompatibility, with 3 of its 277 amino acids being critical for the *Rj2*-mediated transient activation of defense markers such as the *PR-2* gene at two days after inoculation (116). Remarkably, in host plants, a single–amino acid difference between *Rj2* alleles may result in symbiotic incompatibility (117). While NopP of *B. diazoefficiens* USDA110 did not elicit ETI on *G. max* cultivar Hardee (116), it blocked nodulation via root hairs on soybean accessions with a functional *Gm*NNL1 locus, which encodes another TIR-NBS-LRR protein also capable of detecting NopP variants (145). Interestingly, disruption of *Gm*NNL1 by a 179-bp short interspersed nuclear element (SINE)-like transposon improved nodulation via root-hair infection and consequently nitrogen fixation by previously rejected strains. This could explain why today most cultivated soybeans carry an inactivated *Gm*NNL1 locus, which was inadvertently selected by breeders when they screened for improved nitrogen fixation and biomass accumulation (145).

In *G. max* cultivar BARC-2 of the *Rj4/Rj4* genotype, *B. elkanii* strain USDA61 failed to form root-hair curling and infection threads, while roots accumulated salicylic acid, H_2O_2, and products of defense-related genes, some of which were shown to act during ETI in *Arabidopsis* (143). However, the *Rj4* dominant allele of soybean chromosome 1 that restricts nodulation by USDA61 does not code for an NBS-LRR but rather for a thaumatin-like protein (120) that mediates symbiotic incompatibility when challenged by the SUMO-protease effector Bel2-5 (27).

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

T3E-mediated symbiotic incompatibility is not restricted to *Glycine soja* and *G. max* cultivars, as *V. radiata* accessions KPS1 and SUT1 were shown to be incompatible with the *B. elkanii* USDA61 and *B. vignae* ORS3257 strains, respectively. Interestingly, two distinct effectors, NopP of strain ORS3257 and InnB of strain USDA61, were shown to mediate this incompatibility in *V. radiata* (80, 108). The identification of Rj2 orthologs in *V. radiata* suggests that the mechanism responsible for NopP recognition by the soybean Rj2 system may also mediate strain compatibility in *Vigna* spp. As additional T3Es such as InnB and Bel2-5 can also provoke symbiotic incompatibility via recognition by distinct R-proteins, a diverse range of controls appear to be used by legumes to select specific rhizobia.

That legumes may block nodulation of proficient rhizobia raises the question of what can lead to such a deadlock in an interaction that would otherwise be beneficial to the host. One possible explanation is collateral damage caused by the immune system programed to repel microbial infections. Some studies aiming at large-scale identification of effector targets have revealed the existence of a core hub of plant proteins likely to be targeted by a wide range of phytopathogens (76, 137). It is tempting to speculate that some T3Es secreted by rhizobia may also target plant proteins belonging to this so-called immune hub to promote symbiotic infections. If so, some of the NBS-LRR proteins that guard the immune hub against pathogen effectors could recognize rhizobial T3Es and trigger a response, thereby limiting infection of proficient rhizobia.

T3SS ALTERNATIVES TO NOD FACTOR–DEPENDENT NODULATION

The long-established paradigm of nodulation depending exclusively on NFs was challenged when it was shown that photosynthetic *Bradyrhizobium* strains forming nodules on both roots and stems of tropical *Aeschynomene* species did not possess the canonical *nod* genes required for NF synthesis (35). Initially considered an exception to the rule, nodulation in the absence of NFs was recognized as being less exceptional when legume species other than *Aeschynomene* were found to form nodules when inoculated with rhizobia mutants unable to synthesize NFs (68, 84). Remarkably, nodulation of *G. max* cultivar Enrei by a *nodC* mutant of the nonphotosynthetic *B. elkanii* USDA61 strain was shown to depend upon a functional T3SS (84). As several symbionts of *Aeschynomene* species are photosynthetic bradyrhizobia that do not produce NFs and do not possess T3SSs (e.g., strains ORS278 and BTAi1), it now appears that at least two alternative NF-independent pathways may support nodule formation, one of which involves T3Es (85). A feature shared by these NF-independent nodulation processes is that infection of legume tissues by rhizobia is not mediated via infection threads, which is the case in most legume crops. Instead, rhizobia use an intercellular invasion mechanism that still occurs in approximately 25% of known legume species and that is suggested to predate infection threads (12, 84, 113).

Apparently, the use of T3Es to nodulate legumes is widespread in bradyrhizobia, as a wide range of nonphotosynthetic *Bradyrhizobium* strains can elicit nodules on *A. indica* in a T3SS-dependent manner (85). Interestingly, depending on the inoculated strain, infection of nodule cells is not necessarily achieved. For *B. elkanii* strain USDA61, proliferation is restricted to the intercellular space between nodule cells, whereas *B. vignae* strain ORS3257 infects nodule cells intracellularly (**Figure 3**). Although differences in the composition of cocktails of secreted T3Es may contribute to these distinct phenotypes, intracellular ORS3257 bacteria did not fix enough nitrogen to foster plant growth, possibly because bacteroid differentiation was blocked and/or premature degradation of intracellular bacteria occurred (85). In the case of strain ORS3257, at least five T3Es were found to play synergistic and complementary roles in nodulation of *A. indica*: (*a*) A rhizobial-specific T3E named ErnA (effector required for nodulation A) controlled nodule organogenesis, (*b*) NopT and NopAB were both needed for infection of nodules, and (*c*) NopM1 and NopP1 were required to maintain chronic infection of nodule cells (121). However, whether these successive

T3SS-dependent steps in nodulation and infection suffice to secure proficient symbioses in natural environments remains to be demonstrated. In this respect, so far only a few strains such as *Bradyrhizobium mercantei* SEMIA6399, *Bradyrhizobium liaoningense* CCBAU83689, and *Bradyrhizobium* sp. Y36, all of which were isolated from field nodules, have been confirmed to lack *nod* genes and to carry T3SS loci. It is therefore possible that in those cases, T3Es act as the main nodulation determinants (122).

RHIZOBIAL T3Es CAPABLE OF TRIGGERING NODULATION

The importance of ErnA in nodulation of legumes was confirmed using two approaches (121). First, the transfer of *ernA* was shown to extend the host range of a recipient *Bradyrhizobium* strain to *A. indica*. Second, transgenic roots of *A. indica* that expressed *ernA* ectopically formed nodule-like structures (**Figure 3**). The 370-amino-acid-long ErnA displays no homology to known functional domains, except for a nuclear localization signal (NLS) needed for targeting the plant nucleus. As Förster resonance energy transfer (FRET)-fluorescence lifetime imaging (FLIM) data confirmed the association between ErnA and nuclear nucleic acids *in planta*, it was proposed that ErnA modulates plant gene expression and in this way activates the nodule developmental program. In support of this hypothesis, several T3Es from plant pathogens were shown to modulate plant gene expression (15), including the transcription activator–like (TAL) effectors from *Xanthomonas* spp. (e.g., members of the AvrB family) and the two paralogous T3Es from *Pantoea agglomerans* called HsvG and HsvB (7, 8, 15, 81, 82). Interestingly, HsvG and HsvB were found to induce the formation of galls that, like in root nodules, involve the reactivation of the cell cycle in the targeted plant cells.

However, ErnA of strain ORS3257 is not the only rhizobial T3E capable of triggering nodule formation. Notably, while the 1,328-amino-acid-long Bel2-5 of *B. elkanii* USDA61 blocks symbiosis with *G. max* cultivar BARC-2, it was shown to be required for nodulation of the *G. max* cultivar Enrei *nfr1* mutant that is unable to perceive NFs (93). Since Bel2-5 targets plant cell nuclei and carries a SUMO protease domain needed for nodulation (93), Bel2-5 might alter gene expression in legume cells by deSUMOylation of some plant proteins, in much the same way as XopD of *X. campestris* does in *Arabidopsis* and tomato (54, 119). If Bel2-5 and ErnA most probably trigger nodulation using different mechanisms, they both may target conserved determinants in the NF-signaling pathway such as the nuclear regulators CCaMK, CYCLOPS, NIN, LBD16, NF-Y, and CRE1, which, when overexpressed in various legumes, induce nodule-like structures similar to those triggered by ErnA on *A. indica* (26, 36, 101, 106, 111, 112, 123, 124). Alternatively, ErnA and/or Bel2-5 could alter the levels of cytokinin and auxin phytohormones, both of which contribute to nodule formation (9). The fact that T3E HopQ1 of *P. aeruginosa* was shown to activate cytokinin signaling supports this hypothesis (41).

Unlike Bel2-5, which has been found in only a few strains of *B. diazoefficiens*, *B. elkanii*, and *B. japonicum* (93), *ernA* homologs were identified in the genomes of more than two-thirds of the *Bradyrhizobium* strains that carry a T3SS (121, 122). Surprisingly, all of the strains that carried *ernA* and/or *bel2-5* also carried genes for NF synthesis and were isolated from legumes known to depend upon compatible NFs for nodulation. This raises the question of the roles of ErnA and Bel2-5 in conventional NF-dependent symbioses and whether these proteins act synergistically with NFs to reinforce the nodulation properties of a strain. Alternatively, ErnA and Bel2-5 may have been recruited by rhizobia to gain access to legumes that do not respond to the NFs they secrete.

CONCLUSIONS

Initially thought to be characteristic features of pathogens needed for full virulence, T3SSs are now also recognized as hallmarks of beneficial microbes such as rhizobia. In addition to the fact

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

that the T3SS machineries of plant pathogens and rhizobia share structural and functional homologies, symbiotic and pathogenic T3Es act in surprisingly similar ways on the plant cells they target. The fact that T3SS clusters were found to have coevolved with common nodulation genes in most bradyrhizobia (122) suggests that the switch from a free-living to a symbiotic state originally depended on the co-acquisition of T3SSs and NF-biosynthetic functions. Alternatively, in the context of primal rhizobium-legume interactions, the activity of T3SSs was under strong positive selection, at least in members of the *Bradyrhizobium* genus. Either way, and perhaps except for ErnA and Bel2-5, which are capable of triggering nodulation even in the absence of NFs, symbiotic T3Es are mostly involved in specifying the host range of rhizobia, much in the same way as effectors do for plant and animal pathogens. When translocated into plant cells not equipped for their detection, T3Es of rhizobia appear to promote colonization of plant tissues mostly by interfering with the plant systems needed to locally fight and repel microbial infections (e.g., MAPKs in PTI signaling). By contrast, in legume hosts capable of detecting one or several of the secreted T3Es, plant immune responses may halt nodulation. That cocktails of T3Es eventually determine the symbiotic outcome is thus fully consistent with a multifactorial checks-and-balance system that allows plants to carefully choose the bacteria that will be allowed access to the secluded ecological niche formed by nodules.

The lowering of plant defenses below a threshold that would block infection by rhizobia appears to be a common phenomenon in many—if not most—legume-rhizobia associations. Since the infectious perimeter to which symbiotic rhizobia are confined is strikingly circumscribed in both time and space (e.g., infection threads, infection pockets, and some nodule cells), the control of legumes over rhizobium penetration and dissemination is extremely tight. In plant-pathogen research, the continuous evolution of T3E inventories needed for successful infections versus the associated plant defense mechanisms needed to repel pathogens has often been compared to an arms race. Despite the obvious growth benefits legumes obtain by associating with nitrogen-fixing bacteria, host plants and rhizobia also appear to be engaged in a struggle in which bacteria search to secure all-inclusive accommodations while legumes try to protect resources and to maximize nodule outputs. Both ErnA and Bel2-5 may thus be regarded as rhizobial attempts to force the issue toward nodulation, possibly together with other T3Es whose role in triggering nodule organogenesis may have thus far been overshadowed by the more effective NFs.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The work on this review was supported by grants from the French National Research Agency ("SymEffectors"; ANR-16-CE20-0013 and "ET-Nod"; ANR-20-CE20-0012) and from the Swiss National Science Foundation (31003A-173191). A.T. and A.C. were supported by a PhD fellowship from the French Ministry of National Education, Higher Education and Research.

LITERATURE CITED

1. Abby SS, Rocha EPC. 2012. The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. *PLOS Genet*. 8:e1002983
2. Bartsev AV, Deakin WJ, Boukli NM, McAlvin CB, Stacey G, et al. 2004. NopL, an effector protein of *Rhizobium* sp. NGR234, thwarts activation of plant defense reactions. *Plant Physiol*. 134(2):871–79

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

3. Bastedo DP, Lo T, Laflamme B, Desveaux D, Guttman DS. 2020. Diversity and evolution of type III secreted effectors: a case study of three families. *Curr. Top. Microbiol. Immunol.* 427:201–30
4. Bergeron JRC, Fernández L, Wasney GA, Vuckovic M, Reffuveille F, et al. 2016. The structure of a type 3 secretion system (T3SS) ruler protein suggests a molecular mechanism for needle length sensing. *J. Biol. Chem.* 291(4):1676–91
5. Berndsen CE, Wolberger C. 2014. New insights into ubiquitin E3 ligase mechanism. *Nat. Struct. Mol. Biol.* 21(4):301–7
6. Berrabah F, Ratet P, Gourion B. 2019. Legume nodules: massive infection in the absence of defense induction. *Mol. Plant Microbe Interact.* 32:35–44
7. Boch J, Bonas U. 2010. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* 48:419–36
8. Boch J, Bonas U, Lahaye T. 2014. TAL effectors—pathogen strategies and plant resistance engineering. *New. Phytol.* 204(4):823–32
9. Boivin S, Fonouni-Farde C, Frugier F. 2016. How auxin and cytokinin phytohormones modulate root microbe interactions. *Front. Plant Sci.* 7:1240
10. Boller T, Felix G. 2009. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptor. *Annu. Rev. Plant Biol.* 60:379–406
11. Bolzan de Campos S, Deakin WJ, Broughton WJ, Passaglia LMP. 2011. Roles of flavonoids and the transcriptional regulator TtsI in the activation of the type III secretion system of *Bradyrhizobium elkanii* SEMIA587. *Microbiology* 157(3):627–35
12. Bonaldi K, Gargani D, Prin Y, Fardoux J, Gully D, 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(11):1359–71
13. Busset N, Gully D, Teulet A, Fardoux J, Camuel A. et al. 2021. The Type III effectome of the symbiotic *Bradyrhizobium vignae* strain ORS3257. *Biomolecules* 11(11):1592
14. Büttner D. 2012. Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiol. Mol. Biol. Rev.* 76(2):262–310
15. Büttner D. 2016. Behind the lines—actions of bacterial type III effector proteins in plant cells. *FEMS Microbiol. Rev.* 40(6):894–937
16. Caldwell BE. 1966. Inheritance of a strain specific ineffective nodulation in soybeans. *Crop. Sci.* 6:427–28
17. Cascales E, Christie PJ. 2003. The versatile bacterial type IV secretion systems. *Nat. Rev. Microbiol.* 1(2):137–49
18. Cui H, Tsuda K, Parker JE. 2015. Effector-triggered immunity: from pathogen perception to robust defense. *Annu. Rev. Plant Biol.* 66:487–511
19. Dai W-J, Zeng Y, Xie Z-P, Staehelin C. 2008. Symbiosis-promoting and deleterious effects of NopT, a novel type 3 effector of *Rhizobium* sp. strain NGR234. *J. Bacteriol.* 190(14):5101–10
20. de Lajudie PM, Andrews M, Ardley J, Eardly B, Jumas-Bilak E, et al. 2019. Minimal standards for the description of new genera and species of rhizobia and agrobacteria. *Int J. Syst. Evol. Microbiol.* 69(7):1852–63
21. Deakin WJ, Broughton WJ. 2009. Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat. Rev. Microbiol.* 7(4):312–20
22. Deakin WJ, Marie C, Saad MM, Krishnan HB, Broughton WJ. 2005. NopA is associated with cell surface appendages produced by the type III secretion system of *Rhizobium* sp. strain NGR234. *Mol. Plant Microbe Interact.* 18(5):499–507
23. Dean P. 2011. Functional domains and motifs of bacterial type III effector proteins and their roles in infection. *FEMS Microbiol. Rev.* 35(6):1100–25
24. Deng W, Marshall NC, Rowland JL, McCoy JM, Worrall LJ, et al. 2017. Assembly, structure, function and regulation of type III secretion systems. *Nat. Rev. Microbiol.* 15(6):323–37
25. Dowen RH, Engel JL, Shao F, Ecker JR, Dixon JE. 2009. A family of bacterial cysteine protease type III effectors utilizes acylation-dependent and -independent strategies to localize to plasma membranes. *J. Biol. Chem.* 284(23):15867–79
26. Fabre S, Gully D, Poitout A, Patrel D, Arrighi JF, et al. 2015. Nod factor-independent nodulation in *Aeschynomene evenia* required the common plant-microbe symbiotic toolkit. *Plant Physiol.* 169(4):2654–64

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

27. Faruque OM, Miwa H, Yasuda M, Fujii Y, Kaneko T, et al. 2015. Identification of *Bradyrhizobium elkanii* genes involved in incompatibility with soybean plants carrying the *Rj4* allele. *Appl. Environ. Microbiol.* 81(19):6710–17
28. Feng F, Sun J, Radhakrishnan GV, Lee T, Bozsóki Z, et al. 2019. A combination of chitooligosaccharide and lipochitooligosaccharide recognition promotes arbuscular mycorrhizal associations in *Medicago truncatula*. *Nat. Commun.* 10(1):5047
29. Fotiadis CT, Dimou M, Georgakopoulos DG, Katinakis P, Tampakaki AP. 2011. Functional characterization of NopT1 and NopT2, two type III effectors of *Bradyrhizobium japonicum*. *FEMS Microbiol. Lett.* 327(1):66–77
30. Foyer CH, Lam HM, Nguyen HT, Siddique KH, Varshney RK, et al. 2016. Neglecting legumes has compromised human health and sustainable food production. *Nat. Plants* 2:16112
31. Galán JE, Lara-Tejero M, Marlovits TC, Wagner S. 2014. Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells. *Annu. Rev. Microbiol.* 68:415–38
32. Gazi AD, Sarris PF, Fadouloglou VE, Charova SN, Mathioudakis N, et al. 2012. Phylogenetic analysis of a gene cluster encoding an additional, rhizobial-like type III secretion system that is narrowly distributed among *Pseudomonas syringae* strains. *BMC Microbiol.* 12:188
33. Ge YY, Xiang QW, Wagner C, Zhang D, Xie ZP, Staehelin C. 2016. The type 3 effector NopL of *Sinorhizobium* sp. strain NGR234 is a mitogen-activated protein kinase substrate. *J. Exp. Bot.* 67(8):2483–94
34. Gibson KE, Kobayashi H, Walker GC. 2008. Molecular determinants of a symbiotic chronic infection. *Annu. Rev. Genet.* 42:413–41
35. Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E, et al. 2007. Legumes symbioses: absence of *nod* genes in photosynthetic bradyrhizobia. *Science* 316(5829):1307–12
36. Gleason C, Chaudhuri S, Yang T, Muñoz A, Poovaiah BW, Oldroyd GED. 2006. Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature* 441(7097):1149–52
37. Gómez-Gómez L, Felix G, Boller T. 1999. A single locus determines sensitivity to bacterial flagellin in *Arabidopsis thaliana*. *Plant J.* 18(3):277–84a
38. Gottig N, Vranych CV, Sgro GG, Piazza A, Ottado J. 2018. HrpE, the major component of the *Xanthomonas* type three protein secretion pilus, elicits plant immunity responses. *Sci. Rep.* 8:9842
39. Gourion B, Berrabah F, Ratet P, Stacey G. 2015. Rhizobium-legume symbioses: the crucial role of plant immunity. *Trends Plant Sci.* 20:186–94
40. Grohmann E, Christie PJ, Waksman G, Backert S. 2018. Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol. Microbiol.* 107(4):455–71
41. Hann DR, Domínguez-Ferreras A, Motyka V, Dobrev PI, Schornack S, et al. 2014. The *Pseudomonas* type III effector HopQ1 activates cytokinin signaling and interferes with plant innate immunity. *New Phytol.* 201(2):585–98
42. Harris JM, Balint-Kurti P, Bede JC, Day B, Gold S, et al. 2020. What are the Top 10 unanswered questions in molecular plant-microbe interactions? *Mol. Plant Microbe Interact.* 33(12):1354–65
43. Hempel J, Zehner S, Götffert M, Patschkowski T. 2009. Analysis of the secretome of the soybean symbiont *Bradyrhizobium japonicum*. *J. Biotechnol.* 140(1–2):51–58
44. Hotson A, Chosed R, Shu H, Orth K, Mudgett MB. 2003. *Xanthomonas* type III effector XopD targets SUMO-conjugated proteins *in planta*. *Mol. Microbiol.* 50(2):377–89
45. Hubber A, Vergunst AC, Sullivan JT, Hooykaas PJJ, Ronson CW. 2004. Symbiotic phenotypes and translocated effector proteins of the *Mesorhizobium loti* strain R7A VirB/D4 type IV secretion system. *Mol. Microbiol.* 54(2):561–74
46. Hungria M, Menna P, Delamuta JRM. 2015. *Bradyrhizobium*, the ancestor of all rhizobia: phylogeny of housekeeping and nitrogen-fixation genes. In *Biological Nitrogen Fixation*, Vol. 2, ed. F de Bruijn, pp. 191–202. Hoboken, NJ: John Wiley
47. Janczarek M, Kamila Rachwał K, Marzec A, Grządziel J, Palusińska-Szysz M. 2015. Signal molecules and cell-surface components involved in early stages of the legume-rhizobium interactions. *Appl. Soil Ecol.* 85:94–113

48. Jiménez-Guerrero I, Acosta-Jurado S, Medina C, Ollero FJ, Alias-Villegas C, et al. 2020. The *Sinorhizobium fredii* HH103 type III secretion system effector NopC blocks nodulation with *Lotus japonicus* Gifu. *J. Exp. Bot.* 71(19):6043–56
49. Jiménez-Guerrero I, Pérez-Montaño F, Medina C, Ollero FJ, López-Baena FJ. 2015. NopC is a rhizobium-specific type 3 secretion system effector secreted by *Sinorhizobium* (*Ensifer*) *fredii* HH103. *PLOS ONE* 10(11):e0142866
50. Jiménez-Guerrero I, Pérez-Montaño F, Medina C, Ollero FJ, López-Baena FJ. 2017. The *Sinorhizobium* (*Ensifer*) *fredii* HH103 nodulation outer protein NopI is a determinant for efficient nodulation of soybean and cowpea plants. *Appl. Environ. Microbiol.* 83:e02770-16
51. Jones JD, Dangl JL. 2006. The plant immune system. *Nature* 444:323–29
52. Kambara K, Ardissone S, Kobayashi H, Saad MM, Schumpp O, et al. 2009. Rhizobia utilize pathogen-like effector proteins during symbiosis. *Mol. Microbiol.* 71(1):92–106
53. Kim JG, Stork W, Mudgett MB. 2013. *Xanthomonas* type III effector XopD desumoylates tomato transcription factor SlERF4 to suppress ethylene responses and promote pathogen growth. *Cell Host Microbe* 13(2):143–54
54. Kim JG, Taylor, Mudgett MB. 2011. Comparative analysis of the XopD type III secretion (T3S) effector family in plant pathogenic bacteria. *Mol. Plant Pathol.* 12(8):715–30
55. Kimbrel JA, Thomas WJ, Jiang Y, Creason AL, Thireault CA, et al. 2013. Mutualistic co-evolution of type III effector genes in *Sinorhizobium fredii* and *Bradyrhizobium japonicum*. *PLOS Pathog*. 9(2):e1003204
56. Kouchi H, Shimomura K, Hata S, Hirota A, Wu GJ, et al. 2004. Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*. *DNA Res*. 11:263–74
57. Krause A, Doerfel A, Göttfert M. 2002. Mutational and transcriptional analysis of the type III secretion system of *Bradyrhizobium japonicum*. *Mol. Plant Microbe Interact.* 15(12):1228–35
58. Krishnan HB, Lorio J, Kim WS, Jiang G, Kim KY, et al. 2003. Extracellular proteins involved in soybean cultivar-specific nodulation are associated with pilus-like surface appendages and exported by a type III protein secretion system in *Sinorhizobium fredii* USDA257. *Mol. Plant Microbe Interact.* 16(7):617–25
59. Krishnan HB, Natarajan SS, Kim WS. 2011. Distinct cell surface appendages produced by *Sinorhizobium fredii* USDA257 and *S. fredii* USDA191, cultivar-specific and nonspecific symbionts of soybean. *Appl. Environ. Microbiol.* 77(17):6240–48
60. Kusakabe S, Higasitani N, Kaneko T, Yasuda, Miwa H, et al. 2020. *Lotus* accessions possess multiple checkpoints triggered by different type III secretion system effectors of the wide-host-range symbiont *Bradyrhizobium elkanii* USDA61. *Microbes Environ*. 35(1):ME19141
61. Liang Y, Cao Y, Tanaka K, Thibivilliers S, Wan J, et al. 2013. Nonlegumes respond to rhizobial Nod factors by suppressing the innate immune response. *Science* 341(6152):1384–87
62. Liu D, Luo Y, Zheng X, Wang X, Chou M, Wei G. 2021. TRAPPC13 is a novel target of *Mesorhizobium amorphae* type III secretion system effector NopP. *Mol. Plant Microbe Interact.* 34(5):511–23
63. Lohar DP, Sharopova N, Endre G, Penuela S, Samac D, et al. 2006. Transcript analysis of early nodulation events in *Medicago truncatula*. *Plant Physiol*. 140:221–34
64. Lopez-Baena FJ, Vinardell JM, Pérez-Montaño F, Crespo-Rivas JC, Bellogin RA, et al. 2008. Regulation and symbiotic significance of nodulation outer proteins secretion in *Sinorhizobium fredii* HH103. *Microbiology* 154(Part 6):1825–36
65. Lopez-Gomez M, Sandal N, Stougaard J, Boller T. 2012. Interplay of flg22-induced defence responses and nodulation in *Lotus japonicus*. *J. Exp. Bot.* 63(1):393–401
66. Lorio JC, Kim WS, Krishnan HB. 2004. NopB, a soybean cultivar-specificity protein from *Sinorhizobium fredii* USDA257, is a type III secreted protein. *Mol. Plant Microbe Interact.* 17:1259–68
67. Luo Y, Liu D, Jiao S, Liu S, Wang X, et al. 2020. Identification of *Robinia pseudoacacia* target proteins responsive to *Mesorhizobium amphore* CCNWGS0123 effector protein NopT. *J. Exp. Bot.* 71(22):7347–63
68. Madsen LH, Tirichine L, Jurkiewicz A, Sullivan JT, Heckmann AB, et al. 2010. The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat. Commun.* 1:10
69. Marie C, Broughton WJ, Deakin WJ. 2001. *Rhizobium* type III secretion systems: legume charmers or alarmers? *Curr. Opin. Plant Biol.* 4(4):336–42

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

70. Marie C, Deakin WJ, Ojanen-Reuhs T, Diallo E, Reuhs B, et al. 2004. TtsI, a key regulator of *Rhizobium* species NGR234 is required for type III-dependent protein secretion and synthesis of rhamnose-rich polysaccharides. *Mol. Plant Microbe Interact.* 17(9):958–66
71. Marie C, Deakin WJ, Viprey V, Kopciñska J, Golinowski W, et al. 2003. Characterization of Nops, nodulation outer proteins, secreted via the type III secretion system of NGR234. *Mol. Plant Microbe Interact.* 16(9):743–51
72. Masson-Boivin C, Giraud E, Perret X, Batut J. 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol.* 17(10):458–66
73. Mendes R, Garbeva P, Raaijmakers JM. 2013. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol. Rev.* 37:634–63
74. Miwa H, Okazaki S. 2017. How effectors promote beneficial interactions. *Curr. Opin. Plant Biol.* 38:148–54
75. Morrell R, Sadanandom A. 2019. Dealing with stress: a review of plant SUMO proteases. *Front. Plant Sci.* 10:1122
76. Mukhtar MS, Carvunis AR, Dreze M, Epple P, Steinbrenner J, et al. 2011. Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science* 333(6042):596–601
77. Nelson MS, Chun CL, Sadowsky MJ. 2017. Type IV effector proteins involved in the *Medicago-Sinorhizobium* symbiosis. *Mol. Plant Microbe Interact.* 30(1):28–34
78. Nelson MS, Sadowsky MJ. 2015. Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. *Front. Plant Sci.* 6:491
79. Nguyen HP, Ratu STN, Yasuda M, Göttfert M, Okazaki S. 2018. InnB, a novel type III effector of *Bradyrhizobium elkanii* USDA61, controls symbiosis with *Vigna* species. *Front. Microbiol.* 9:3155
80. Nguyen HP, Ratu STN, Yasuda M, Teaumroong N, Okazaki S. 2020. Identification of *Bradyrhizobium elkanii* USDA61 type III effectors determining symbiosis with *Vigna mungo*. *Genes* 11(5):474
81. Nissan G, Manulis-Sasson S, Chalupowicz L, Teper D, Yeheskel A, et al. 2012. The type III effector HsvG of the gall-forming *Pantoea agglomerans* mediates expression of the host gene *HSVGT*. *Mol. Plant Microbe Interact.* 25(2):231–40
82. Nissan G, Manulis-Sasson S, Weinthal D, Mor H, Sessa G, Barash I. 2006. The type III effectors HsvG and HsvB of gall-forming *Pantoea agglomerans* determine host specificity and function as transcriptional activators. *Mol. Microbiol.* 61(5):1118–31
83. Notti RQ, Stebbins CE. 2016. The structure and function of Type III secretion systems. *Microbiol. Spectr.* 4(1). **https://doi.org/10.1128/microbiolspec.VMBF-0004-2015**
84. Okazaki S, Kaneko T, Sato S, Saeki K. 2013. Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *PNAS* 110(42):17131–36
85. Okazaki S, Tittabutr P, Teulet A, Thouin J, Fardoux J, et al. 2016. Rhizobium–legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. *ISME J.* 10(1):64–74
86. Oldroyd GED. 2013. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* 11(4):252–63
87. Parker MA. 2015. The spread of *Bradyrhizobium* lineages across host legume clades: from *Abarema* to *Zygia*. *Microb. Ecol.* 69(3):630–40
88. Perret X, Staehelin C, Broughton WJ. 2000. Molecular basis of symbiotic promiscuity. *Microbiol. Mol. Biol.* 64(1):180–201
89. Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH. 2013. Going back to the roots: the microbial ecology of the rhizosphere. *Nat. Rev. Microbiol.* 11:789–99
90. Piromyou P, Nguyen N, Songwattana P, Boonchuen P, Teamtisong K, et al. 2021. The *Bradyrhizobium diazoefficiens* type III efector NopE modulates the regulation of plant hormones towards nodulation in *Vigna radiata*. *Sci. Rep.* 11:16604
91. Piromyou P, Songwattana P, Teamtisong K, Tittabutr P, Boonkerd N, et al. 2019. Mutualistic coevolution of T3SSs during the establishment of symbiotic relationships between *Vigna radiata* and bradyrhizobia. *MicrobiologyOpen* 8:e781
92. Poole P, Ramachandran V, Terpolilli J. 2018. Rhizobia: from saprophytes to endosymbionts. *Nat. Rev. Microbiol.* 16(5):291–303

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

93. Ratu STN, Teulet A, Miwa H, Masuda S, Nguyen HP, et al. 2021. Rhizobia use a pathogenic-like effector to hijack leguminous nodulation signaling. *Sci. Rep.* 11(1):2034
94. Remigi P, Zhu J, Young JPW, Masson-Boivin C. 2016. Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. *Trends Microbiol.* 24(1):63–75
95. Reuhs BL, Relić B, Forsberg LS, Marie C, Ojanen-Reuhs T, et al. 2005. Structural characterization of a flavonoid-inducible *Pseudomonas aeruginosa* A-band-like O antigen of *Rhizobium* sp. strain NGR234, required for the formation of nitrogen-fixing nodules. *J. Bacteriol.* 187(18):6479–87
96. Rohde JR, Breitkreutz A, Chenal A, Sansonetti PJ, Parsot C. 2007. Type III secretion effectors of the IpaH family are E3 ubiquitin ligases. *Cell Host Microbe* 1(1):77–83
97. Saad MM, Crèvecoeur M, Masson-Boivin C, Perret X. 2012. The type 3 protein secretion system of *Cupriavidus taiwanensis* strain LMG19424 compromises symbiosis with *Leucaena leucocephala*. *Appl. Environ. Microbiol.* 78(20):7476–79
98. Saad MM, Kobayashi H, Marie C, Brown IR, Mansfield JW, et al. 2005. NopB, a type III secreted protein of *Rhizobium* sp. strain NGR234, is associated with pilus-like surface appendages. *J. Bacteriol.* 187(3):1173–81
99. Saad MM, Staehelin C, Broughton WJ, Deakin WJ. 2008. Protein-protein interactions within type III secretion system-dependent pili of *Rhizobium* sp. strain NGR234. *J Bacteriol.* 190(2):750–54
100. Schechter LM, Guenther J, Olcay EA, Sungchan J, Krishnan HB. 2010. Translocation of NopP by *Sinorhizobium fredii* USDA257 into *Vigna unguiculata* root nodules. *Appl. Environ. Microbiol.* 76(11):3758–61
101. Schiessl K, Lilley JLS, Lee T, Tamvakis I, Kohlen W. 2019. *NODULE INCEPTION* recruits the lateral root developmental program for symbiotic nodule organogenesis in *Medicago truncatula*. *Curr. Biol.* 29(21):3657–68.e5
102. Schirrmeister J, Friedrich L, Wenzel M, Hoppe M, Wolf C, et al. 2011. Characterization of the self-cleaving effector protein NopE1 of *Bradyrhizobium japonicum*. *J. Bacteriol.* 193(15):3733–39
103. Schmeisser C, Liesegang H, Krysciak D, Bakkou N, Le Quéré A, et al. 2009. *Rhizobium* sp. strain NGR234 possesses a remarkable number of secretion systems. *Appl. Environ. Microbiol.* 75(12):4035–45
104. Sharma B, Joshi D, Yadav PK, Gupta AK, Bhatt TK. 2016. Role of ubiquitin-mediated degradation system in plant biology. *Front. Plant Sci.* 7:806
105. Sheikh TMM, Zhang L, Zubair M, Hanif A, Li P. 2019. The type III accessory protein HrpE of *Xanthomonas oryzae* pv. *oryzae* surpasses the secretion role, and enhances plant resistance and photosynthesis. *Microorganisms* 7(11):572
106. Singh S, Katzer K, Lambert J, Cerri M, Parniske M. 2014. CYCLOPS, a DNA-binding transcriptional activator, orchestrates symbiotic root nodule development. *Cell Host Microbe* 15(2):139–52
107. Skorpil P, Saad MM, Boukli NM, Kobayashi H, Ares-Orpel F, et al. 2005. NopP, a phosphorylated effector of *Rhizobium* sp. strain NGR234, is a major determinant of nodulation of the tropical legumes *Flemingia congesta* and *Tephrosia vogelii*. *Mol. Microbiol.* 57(5):1304–17
108. Songwattana P, Chaintreuil C, Wongdee J, Teulet A, Mbaye M, et al. 2021. Identification of type III effectors modulating the symbiotic properties of *Bradyrhizobium vignae* strain ORS3257 with various *Vigna* species. *Sci. Rep.* 11(1):4874
109. Songwattana P, Noisongiam R, Teamtisong K, Prakamhang J, Teulet A, et al. 2017. Type 3 secretion system (T3SS) of *Bradyrhizobium* sp. DOA9 and its roles in legume symbiosis and rice endophytic association. *Front. Microbiol.* 8:1810
110. Songwattana P, Tittabutr P, Wongdee J, Teamtisong K, Wulandari D, et al. 2019. Symbiotic properties of a chimeric Nod-independent photosynthetic *Bradyrhizobium* strain obtained by conjugative transfer of a symbiotic plasmid. *Environ. Microbiol.* 21(9):3442–54
111. Soyano T, Kouchi H, Hirota A, Hayashi M. 2013. Nodule inception directly targets NF-Y subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. *PLOS Genet.* 9(3):e1003352
112. Soyano T, Shimoda Y, Kawaguchi M, Hayashi M. 2019. A shared gene drives lateral root development and root nodule symbiosis pathways in *Lotus*. *Science* 366:1021–23
113. Sprent JI. 2007. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytol.* 174(1):11–25

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

114. Staehelin C, Krishnan HB. 2015. Nodulation outer proteins: double-edged swords of symbiotic rhizobia. *Biochem. J.* 470(3):263–74
115. Sugawara M, Epstein B, Badgley BD, Unno T, Xu L, et al. 2013. Comparative genomics of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. *Genome Biol.* 14(2):R17
116. Sugawara M, Takahashi S, Umehara Y, Iwano H, Tsurumaru H, et al. 2018. Variation in bradyrhizobial NopP effector determines symbiotic incompatibility with *Rj2*-soybeans via effector-triggered immunity. *Nat. Commun.* 9(1):3139
117. Sugawara M, Umehara Y, Kaga A, Hayashi M, Ishimoto M, et al. 2019. Symbiotic incompatibility between soybean and *Bradyrhizobium* arises from one amino acid determinant in soybean Rj2 protein. *PLOS ONE* 14(9):e0222469
118. Tampakaki AP. 2014. Commonalities and differences of T3SSs in rhizobia and plant pathogenic bacteria. *Front. Plant Sci.* 5:114
119. Tan L, Rong W, Luo H, Chen Y, He C. 2014. The *Xanthomonas campestris* effector protein XopDXcc8004 triggers plant disease tolerance by targeting DELLA proteins. *New Phytol.* 204(3):595–608
120. Tang F, Yang S, Liu J, Zhu H. 2016. *Rj4*, a gene controlling nodulation specificity in soybeans, encodes a thaumatin-like protein but not the one previously reported. *Plant Physiol.* 170:26–32
121. Teulet A, Busset N, Fardoux J, Gully D, Chaintreuil C, et al. 2019. The rhizobial type III effector ErnA confers the ability to form nodules in legumes. *PNAS* 116(43):21758–68
122. Teulet A, Gully D, Rouy Z, Camuel A, Koebnik R, et al. 2020. Phylogenetic distribution and evolutionary dynamics of *nod* and T3SS genes in the genus *Bradyrhizobium*. *Microb. Genom.* 6(9):mgen000407
123. Tirichine L, Imaizumi-Anraku H, Yoshida S, Murakami Y, Madsen LH, et al. 2006. Deregulation of a Ca^{2+}/calmodulin-dependent kinase leads to spontaneous nodule development. *Nature* 441(7097):1153–56
124. Tirichine L, Sandal N, Madsen LH, Radutoiu S, Albrektsen AS, et al. 2007. A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science* 315(5808):104–7
125. Tong W, Li X, Wang E, Cao Y, Chen W, et al. 2020. Genomic insight into the origins and evolution of symbiosis genes in *Phaseolus vulgaris* microsymbionts. *BMC Genom.* 21(1):186
126. Troisfontaines P, Cornelis GR. 2005. Type III secretion: more systems than you think. *Physiology* 20:326–39
127. Van Engelenburg SB, Palmer AE. 2008. Quantification of real-time *Salmonella* effector type-III secretion kinetics reveals differential secretion rates for SopE2 and SptP. *Chem. Biol.* 15(6):619–28
128. Vasse J, de Billy F, Truchet G. 1993. Abortion of infection during the *Rhizobium meliloti*-alfalfa symbiotic interaction is accompanied by a hypersensitive reaction. *Plant J.* 4(3):555–66
129. Viprey V, Del Greco A, Golinowski W, Broughton WJ, Perret X. 1998. Symbiotic implications of type III protein secretion machinery in *Rhizobium*. *Mol. Microbiol.* 28(6):1381–89
130. Wagner S, Grin I, Malmsheimer S, Singh N, Torres-Vargas CE, Westerhausen S. 2018. Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into eukaryotic host cells. *FEMS Microbiol Lett.* 365(19):fny201
131. Wang J, Wang J, Liu C, Ma C, Li C, et al. 2018. Identification of soybean genes whose expression is affected by the *Ensifer fredii* HH103 effector protein NopP. *Int. J. Mol. Sci.* 19(11):3438
132. Wang J, Wang J, Ma C, Zhou Z, Yang D, et al. 2020. QTL mapping and data mining to identify genes associated with the *Sinorhizobium fredii* HH103 T3SS effector NopD in soybean. *Front. Plant Sci.* 11:453
133. Wang S, Meade A, Lam HM, Luo H. 2020. Evolutionary timeline and genomic plasticity underlying the lifestyle diversity in Rhizobiales. *mSystems* 5(4):e00438-20
134. Wang W, Feng B, Zhou JM, Tang. 2020. Plant immune signaling: advancing on two frontiers. *J. Integr. Plant Biol.* 62(1):2–24
135. Wassem R, Kobayashi H, Kambara K, Le Quéré A, Walker GC, et al. 2008. TtsI regulates symbiotic genes in *Rhizobium* species NGR234 by binding to *tts* boxes. *Mol. Microbiol.* 68(3):736–48
136. Wenzel M, Friedrich L, Göttfert M, Zehner S. 2010. The type III-secreted protein NopE1 affects symbiosis and exhibits a calcium-dependent autocleavage activity. *Mol. Plant Microbe Interact.* 23(1):124–29
137. Weßling R, Epple P, Altmann S, He Y, Yang L, et al. 2014. Convergent targeting of a common host protein-network by pathogen effectors from three kingdoms of life. *Cell Host Microbe* 16(3):364–75

138. Winnen B, Schlumberger AC, Sturm A, Schüpbach K, Siebenmann S, et al. 2008. Hierarchical effector protein transport by the *Salmonella* Typhimurium SPI-1 type III secretion system. *PLOS ONE* 3(5):e2178
139. Xiang Q-W, Bai J, Cai J, Huang Q-Y, Wang Y, et al. 2020. NopD of *Bradyrhizobium* sp. XS1150 possesses SUMO protease activity. *Front. Microbiol.* 11:386
140. Xin D-W, Liao S, Xie Z-P, Hann DR, Steinle L, et al. 2012. Functional analysis of NopM, a novel E3 ubiquitin ligase (NEL) domain effector of *Rhizobium* sp. strain NGR234. *PLOS Pathog.* 8(5):e1002707
141. Xu C-C, Zhang D, Hann DR, Xie Z-P, Staehelin C. 2018. Biochemical properties and *in planta* effects of NopM, a rhizobial E3 ubiquitin ligase. *J. Biol. Chem.* 93(39):15304–15
142. Yang S, Tang F, Gao M, Krishnan HB, Zhu H. 2010. *R* gene-controlled host specificity in the legume–rhizobia symbiosis. *PNAS* 107:18735–40
143. Yasuda M, Miwa H, Masuda S, Takebayashi Y, Sakakibara H, Okazaki S. 2016. Effector-triggered immunity determines host genotype-specific incompatibility in legume–*Rhizobium* symbiosis. *Plant Cell Physiol.* 57(8):1791–800
144. Zehner S, Schober G, Wenzel M, Lang K, Göttfert M. 2008. Expression of the *Bradyrhizobium japonicum* type III secretion system in legume nodules and analysis of the associated *tts* box promoter. *Mol. Plant Microbe Interact.* 21(8):1087–93
145. Zhang B, Wang M, Sun Y, Zhao P, Liu C, et al. 2021. *Glycine max* NNL1 restricts symbiotic compatibility with widely distributed bradyrhizobia via root hair infection. *Nat. Plants* 7(1):73–86
146. Zhang J, Li W, Xiang T, Liu Z, Laluk K, et al. 2010. Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a *Pseudomonas syringae* effector. *Cell Host Microbe* 7(4):290–301
147. Zhang L, Chen X-J, Lu H-B, Xie Z-P, Staehelin C. 2011. Functional analysis of the type 3 effector nodulation outer protein L (NopL) from *Rhizobium* sp. NGR234: symbiotic effects, phosphorylation, and interference with mitogen-activated protein kinase signaling. *J. Biol. Chem.* 286(37):32178–87
148. Zhang Y, Liu X, Chen L, Fu Y, Li C, et al. 2018. Mining for genes encoding proteins associated with NopL of *Sinorhizobium fredii* HH103 using quantitative trait loci in soybean *Glycine max* (Merr.) recombinant inbred lines. *Plant Soil* 431:245–55

Annu. Rev. Microbiol. 2022.76:45-65. Downloaded from www.annualreviews.org
Access provided by CIRAD on 01/13/23. For personal use only.

7. Présentation du sujet de thèse

La découverte de l'existence d'un processus symbiotique indépendant de la signalisation médiée par les FNs mais nécessitant un T3SS fonctionnel a permis d'élargir notre champ de vision sur le rôle que pouvait jouer ce système de sécrétion initialement identifié chez les bactéries pathogènes. A travers les études réalisées dans l'équipe MSLT[114] du LSTM[115] et par le groupe de Shin Okazaki en 2013 et 2016, il a été montré que ce processus, bien qu'original, semblait être répandu chez une large diversité de *Bradyrhizobium* non-photosynthétiques. Dès lors, il a été vite admis que les T3Es pouvaient déclencher le processus d'organogenèse nodulaire. Dans le but de mieux comprendre la nature des protéines effectrices impliquées dans la symbiose Nod-indépendante, T3SS-dépendante, une mutagénèse quasi-exhaustive des effecteurs prédits chez la souche *B. vignae* ORS3257 ainsi que sur certains retrouvés chez *B. elkanii* USDA61 a révélé l'existence d'une nouvelle classe d'effecteurs déclenchant directement la nodulation, appelée ET-Nod, dont ErnA et Bel2-5 sont les deux premiers représentants. Le début de la caractérisation fonctionnelle d'ErnA réalisée durant la thèse d'Albin Teulet ainsi que celle réalisée sur Bel2-5 suggère que ces effecteurs pourraient déclencher le programme de nodulation en interférant avec l'expression de certains gènes de l'hôte. La découverte de ces deux premiers ET-Nods a ainsi soulevé de nombreuses questions, parmi lesquelles :

- Quelle est la prévalence des *Bradyrhizobium* à utiliser un processus Nod-indépendant, T3SS-dépendant ?
- Existe-t-il d'autres effecteurs distincts d'ErnA et Bel2-5 capables de déclencher la nodulation chez les *Bradyrhizobium* ?
- Quel est le mode d'action des ET-Nods et comment ces effecteurs activent-ils le processus de nodulation une fois dans la cellule hôte ?
- Est-ce que les différents ET-Nods identifiés utilisent les mêmes voies de signalisation pour initier l'organogenèse nodulaire ou agissent-ils différemment les uns des autres ?
- Les déterminants symbiotiques identifiés dans la voie de signalisation Nod-indépendante, T3SS-indépendante sont-ils aussi recrutés durant le processus T3SS-dépendant ?

C'est sur la base de ce questionnement que j'ai mené les différents travaux qui constituent ma thèse. Dans un premier chapitre qui forme la partie majeure de mon travail de thèse, j'ai utilisé l'espèce *A. indica* car elle est strictement nodulée de manière Nod-indépendante (sans interférence avec les FNs

[114] Mécanismes Symbiotiques chez les Légumineuses Tropicales.
[115] Laboratoire des Symbioses Tropicales et Méditerranéennes.

synthétisés par la bactérie, à l'inverse du soja) et a été utilisée dans les précédentes études réalisées sur ce sujet. D'autre part, la diversité disponible grandissante des souches séquencées de *Bradyrhizobium* a poussé l'équipe MSLT, dans laquelle j'ai réalisé ma thèse, à collecter une grande variété de souches provenant du monde entier et isolée de multiples environnements pour mieux déterminer cette capacité symbiotique à noduler *Aeschynomene* de manière T3SS-dépendante. Ainsi, l'une des principales tâches de cet axe a été de réaliser des tests symbiotiques sur *A. indica* à partir d'une collection d'environ 200 souches de *Bradyrhizobium* séquencés[116] et de tester par mutagénèse des nouveaux ET-Nod qui pouvaient émerger par génomique comparative.

Ces travaux ont permis de montrer que le processus T3SS-dépendant est répandu au sein du genre *Bradyrhizobium* puisqu'environ un tiers des souches testées ont été capables d'induire des nodules *via* leur T3SS. L'analyse du contenu en gènes de ces souches a ensuite conduit à l'identification d'un nouvel ET-Nod nommé Sup3[117] et codant pour une SUMO-protéase tout comme Bel2-5. Des premières analyses fonctionnelles ont permis de montrer que Sup3 est adressé au noyau et que son expression ectopique dans des racines d'*Aeschynomene* conduit à la formation de nodules spontanés. Il a aussi été mis en évidence l'existence de trois autres T3Es putatifs nommés Ubi1, Ubi2 et Ubi3[118], qui semblent jouer un rôle synergique dans la capacité à noduler chez la souche *Bradyrhizobium icense* LMTR13. Enfin, l'ensemble de cette étude a permis de montrer une diversité d'ET-Nods qui était insoupçonnée à ce jour.

Dans le deuxième chapitre de ma thèse, je me suis intéressée à comprendre si les deux processus Nod-indépendant possibles chez *Aeschynomene* recrutaient ou non des acteurs communs pour l'activation de la voie de signalisation conduisant à la nodulation. Dans ce sens, j'ai pu mettre à profit l'existence du modèle *A. evenia* pour lequel des mutants EMS dans les principaux gènes symbiotiques requis pour la symbiose Nod-indépendante, T3SS-indépendante sont disponibles. Cependant, la lignée[119] N76 d'*A. evenia* utilisée pour la mutagénèse est très mal nodulée par les *Bradyrhizobium* non-photosynthétique utilisant un processus T3SS-dépendant. Il a donc fallu identifier une autre lignée plus favorable pour mener des analyses fonctionnelles sur les ET-Nods. La lignée N21 a ainsi émergé comme modèle pour l'étude de la symbiose T3SS-dépendante et une première analyse transcriptomique a été réalisée dans le but d'identifier les potentiels gènes sous le contrôle d'ErnA. Néanmoins, l'intérêt de travailler sur la lignée N76 demeure, en raison de l'existence de mutants dans des gènes symbiotiques clés. Ainsi, il a été montré que selon les souches inoculées et le cocktail d'effecteurs sécrétés dans la plante hôte, la capacité à contourner certains acteurs clés de la sym-

[116] Travail que j'avais pu initier durant mon stage de Master 1 en 2019.
[117] SUMO-protease 3.
[118] Unknown *Bradyrhizobium icense* 1/2/3.
[119] Une lignée se définie par une homogénéité du génotype après plusieurs autofécondations successives.

biose est variable mais que dans tous les cas, AeNSP2 et AeNIN sont requis pour activer le programme de nodulation.

En complément de ces travaux, un autre objectif de ma thèse a été de produire le matériel biologique et d'initier les expérimentations afin de caractériser la ou les cibles moléculaires d'ErnA pour mieux comprendre le mode d'action de cet ET-Nod au cours de la symbiose chez *Aeschynomene*. Cette étude présentée dans le chapitre 3 est toujours en cours et se fait en collaboration avec l'équipe de Laurent Deslandes (LIPM, Toulouse, spécialisée dans la caractérisation fonctionnelle d'effecteurs phytopathogènes) et celle de Peter Mergaert (I2BC, Gif-sur-Yvette, spécialisée dans la symbiose rhizobium-légumineuse). Des premiers gènes cibles candidats sous le contrôle d'ErnA ainsi que ses potentiels interactants directs ont émergé et il convient à présent de les valider.

Dernièrement, à la fin de ma deuxième année de thèse, j'ai bénéficié de nouvelles ressources génomiques qui ont émergé au sein du supergroupe des *Bradyrhizobium* photosynthétiques grâce à une nouvelle collaboration avec le groupe de Haiwei Luo (Université de Hong Kong). En effet, plusieurs centaines de nouvelles souches de *Bradyrhizobium* photosynthétiques ont été isolées et séquencées et proviennent principalement de rizières. L'idée a été, comme dans le premier chapitre ma thèse, de coupler l'analyse des propriétés symbiotiques de ces souches avec leur contenu en gènes par comparaison génomique de diverses sous-populations nouvellement identifiées dans ce phylogroupe afin de faire ressortir des déterminants spécifiques de la symbiose Nod-indépendante, T3SS-indépendante. Cependant à notre grande surprise, l'ensemble des souches représentantes de chacune de ces sous populations ont été trouvées capables de noduler *A. indica*. Ce résultat suggère que les gènes symbiotiques clés gouvernant la symbiose Nod-indépendante, T3SS-indépendante pourraient faire partie du core-génome bactérien du phylogroupe des *Bradyrhizobium* photosynthétiques, ce qui ouvre de nouvelles perspectives pour l'identification des déterminants "non Nod".

En parallèle de mon travail de thèse, d'autres aspects secondaires ont été abordés pour étendre notre connaissance sur le rôle du T3SS et de ses T3Es associés durant l'interaction avec des légumineuses d'intérêt agronomique. Ainsi, des études phylogénétiques sur les gènes *nod* et ceux du T3SS ont permis de mieux comprendre l'évolution de ces gènes et de leur importance dans le genre *Bradyrhizobium*. De plus, au cours d'une collaboration avec l'équipe de Neung Teaumroong (Université de Suranaree, Thaïlande), j'ai participé à l'identification des T3Es modulant l'activité symbiotique de *B. vignae* ORS3257 sur différentes espèces de *Vigna* spp. cultivées en Afrique et en Asie. Enfin, l'effectome complet de *B. vignae* ORS3257 a été obtenu et validé par des approches multi-omics avec le groupe de Peter Mergaert. L'ensemble de ces études complémentaires est disponible en annexe 1.

Chapitre 1

1. Contexte de l'étude

Pour mieux comprendre la diversité des mécanismes symbiotiques chez les souches de *Bradyrhizobium*, l'équipe MSLT dans laquelle j'ai effectué ma thèse a collecté pendant plus de 3 ans des souches dont la séquence du génome était disponible. Ces *Bradyrhizobium* proviennent du monde entier et ont été isolés à partir de matériels biologiques variés (nodules, feuilles, rhizosphère, eaux usées, etc.). L'établissement de cette collection, qui s'élève à presque 200 souches, représente une bonne partie de la diversité disponible et contient entre autres plus de 60 souches types de *Bradyrhizobium*. Elle a pour but de caractériser la capacité des *Bradyrhizobium* à utiliser un processus alternatif indépendant des FNs pour déclencher la nodulation et d'identifier les principaux déterminants génétiques impliqués. En effet, en effectuant des tests symbiotiques sur *A. indica* (groupe 3 d'inoculation croisée), on s'affranchit des FNs et des effets qu'ils peuvent avoir chez les souches en possédant. Par ailleurs, la découverte en 2019 d'ErnA puis en 2021 de Bel2-5 chez *B. vignae* ORS3257 et *B. elkanii* USDA61 respectivement, laisse supposer que d'autres effecteurs pourraient déclencher le programme de nodulation au sein de la diversité du genre *Bradyrhizobium*.

L'objectif principal de mes travaux de thèse a alors été d'étudier la prévalence des *Bradyrhizobium* à utiliser un processus symbiotique T3SS-dépendant, d'identifier de nouveaux ET-Nod capables de déclencher le processus d'organogenèse nodulaire ainsi que d'initier leur caractérisation fonctionnelle. Pour cela, 196 souches ont été testées au moins deux fois sur *A. indica* afin d'identifier les propriétés symbiotiques de chacune d'entre-elles. Il est apparu que plus d'un tiers des souches (35 %) testées sont capables de noduler ou d'induire la formation de pseudo-nodules. En parallèle, la disponibilité du génome entier séquencé de ces souches a permis la construction d'un arbre phylogénomique afin d'essayer de corréler les propriétés symbiotiques des souches testées et leur appartenance à certains clades phylogénétiques. A l'aide de ce support et à travers des analyses génomiques comparatives, la recherche de nouveaux ET-Nods a pu être menée. La méthodologie utilisée s'appuie sur l'obtention et la comparaison des effectomes putatifs des souches ne nodulant pas *A. indica* avec celles capables de noduler mais ne possédant pas d'homologues d'ErnA. C'est au total 14 souches sur les 196 testées qui nodulent et qui n'ont pas ErnA, celles-ci appartenant au supergroupe *B. elkanii* et *B. jacamae*. L'obtention de leur catalogue de T3Es respectifs a été réalisée selon la méthode précédemment utilisée par l'équipe (Teulet et al., 2019). En s'appuyant sur ces données, nous avons pu, par mutagénèse des principaux gènes candidats, identifier un premier nouvel ET-Nod chez la souche *Bradyrhizobium* sp. NAS96.2 (*B. elkanii* supergroupe) codant pour une putative SUMO-protéase : Sup3.

De manière intéressante, cet effecteur est retrouvé chez quatre autres souches capables de noduler *A. indica* alors qu'elles ne possèdent pas ErnA, telle que *B. archetypum* WSM1744 (*B. jicamae* supergroupe). Nous avons pu montrer chez cette dernière souche que Sup3 est également responsable du phénotype de nodulation. La sécrétion de Sup3 par le T3SS a été démontré par l'équipe de Laurent Deslandes (LIPME, Toulouse), ce qui permet de confirmer que Sup3 est bien un T3E. Tout comme ErnA, il a été montré que cet effecteur cible le noyau de la plante hôte, laissant penser que Sup3 pourrait agir directement ou indirectement sur l'expression de certains gènes cibles pour activer le programme de nodulation. Enfin, nous avons pu montrer que l'expression ectopique du gène *sup3* dans des racines transformées d'*A. evenia* ou *A. indica* conduit à la formation de pseudo-nodules spontanés en absence de la bactérie confirmant que Sup3 est bien un ET-Nod.

Chez les souches de *Bradyrhizobium* nodulantes *A. indica* ne possédant ni ErnA ni Sup3, d'autre(s) ET-Nod(s) reste(nt) à identifier. Après comparaison des effectomes putatifs de ces souches qui appartiennent pour la plupart au supergroupe *B. jicamae*, aucun candidat évident ne nous est apparu. C'est pourquoi nous avons décidé de nous focaliser sur la souche *B. icense* LMTR13 : celle-ci possède un catalogue d'effecteurs réduit comparé aux souches voisines. C'est ainsi 10 mutants qui ont été testés. Cependant, aucun ne perd totalement sa capacité à noduler. En revanche, trois mutants Ω*ubi1*, Ω*ubi2* et Ω*ubi3* présentent une réduction significative du nombre de nodules induits sur *A. indica*. Ces observations pourraient indiquer une redondance fonctionnelle entre ces trois nouveaux effecteurs, qui pourraient activer de manière synergique le programme de nodulation. Pour corroborer que ces trois nouveaux effecteurs pourraient correspondre à des ET-Nods, nous avons montré qu'ils partagent des domaines communs à ErnA, ce qui suggère un lien fonctionnel entre ces différents T3Es. Par mutagénèse dirigée, nous avons pu démontrer l'importance de ces domaines dans la fonctionnalité d'ErnA. Il pourrait être intéressant par la suite d'avancer dans la caractérisation fonctionnelle de ces trois nouveaux ET-Nods émergents et pour lesquels aucune étude n'a été réalisée sur des homologues présents chez d'autres souches.

L'ensemble de cette étude a permis de découvrir de nouveaux ET-Nods, parmi lesquels la famille des SUMO-protéases semble jouer un rôle important. De plus, ces travaux nous ont permis d'avoir une vue plus globale sur la diversité des souches de *Bradyrhizobium* capables de noduler de manière indépendante des FNs mais dépendante du T3SS.

2. Article 1 : "Widespread *Bradyrhizobium* distribution of diverse Type III effectors that trigger legume nodulation in the absence of Nod factor"

www.nature.com/ismej

ISME

ARTICLE OPEN

Check for updates

Widespread *Bradyrhizobium* distribution of diverse Type III effectors that trigger legume nodulation in the absence of Nod factor

Alicia Camuel[1,2], Albin Teulet[1,3], Mélanie Carcagno[4], Fazal Haq[5], Valérie Pacquit[4], Djamel Gully[1,2], Marjorie Pervent[1,2], Clémence Chaintreuil[1,2], Joël Fardoux[1], Natasha Horta-Araujo[1,2], Shin Okazaki[6], Safirah Tasa Nerves Ratu[6], Fatou Gueye[7], Jerri Zilli[8], Nico Nouwen[1,2], Jean-François Arrighi[1,2], Haiwei Luo[9], Peter Mergaert[5], Laurent Deslandes[4] and Eric Giraud[1,2]✉

© The Author(s) 2023

The establishment of the rhizobium-legume symbiosis is generally based on plant perception of Nod factors (NFs) synthesized by the bacteria. However, some *Bradyrhizobium* strains can nodulate certain legume species, such as *Aeschynomene* spp. or *Glycine max*, independently of NFs, and via two different processes that are distinguished by the necessity or not of a type III secretion system (T3SS). ErnA is the first known type III effector (T3E) triggering nodulation in *Aeschynomene indica*. In this study, a collection of 196 sequenced *Bradyrhizobium* strains was tested on *A. indica*. Only strains belonging to the photosynthetic supergroup can develop a NF-T3SS-independent symbiosis, while the ability to use a T3SS-dependent process is found in multiple supergroups. Of these, 14 strains lacking *ernA* were tested by mutagenesis to identify new T3Es triggering nodulation. We discovered a novel T3E, Sup3, a putative SUMO-protease without similarity to ErnA. Its mutation in *Bradyrhizobium* strains NAS96.2 and WSM1744 abolishes nodulation and its introduction in an *ernA* mutant of strain ORS3257 restores nodulation. Moreover, ectopic expression of *sup3* in *A. indica* roots led to the formation of spontaneous nodules. We also report three other new T3Es, Ubi1, Ubi2 and Ubi3, which each contribute to the nodulation capacity of strain LMTR13. These T3Es have no homology to known proteins but share with ErnA three motifs necessary for ErnA activity. Together, our results highlight an unsuspected distribution and diversity of T3Es within the *Bradyrhizobium* genus that may contribute to their symbiotic efficiency by participating in triggering legume nodulation.

The ISME Journal (2023) 17:1416–1429; https://doi.org/10.1038/s41396-023-01458-1

INTRODUCTION

Nature has the remarkable capacity to evolve different pathways to reach the same phenotype. This is the case for the signaling cascade initiating the rhizobium-legume symbiosis, which leads to the formation of a new plant organ, the nodule, where the bacteria fix nitrogen for the plant's benefit and receive in exchange photosynthates and other nutrients. The most common way to achieve the formation of this symbiotic organ involves flavonoids on the plant side and Nod factors (NFs) on the rhizobial side [1, 2]. This NF-dependent pathway has long been considered as a universal paradigm governing all rhizobium-legume symbioses. However, two other alternative pathways have been described during the last two decades [3–5]. The first one has been observed in some tropical legume species of the *Aeschynomene* genus that are nodulated by photosynthetic bradyrhizobia lacking the canonical *nod* genes required for NF synthesis [3]. The second was initially identified in *Glycine max*, a legume species known to normally require NFs to initiate symbiosis [4]. It has been shown that the cultivar Enrei can still be nodulated by a *nod* mutant of the non-photosynthetic strain *Bradyrhizobium elkanii* USDA61 [4]. This second alternative process differs to the first one in that this interaction relies on the functioning of a type III secretion system (T3SS), such secretory machinery being absent in most photosynthetic bradyrhizobia nodulating *Aeschynomene* [6]. In laboratory conditions, strain USDA61 as well as a large diversity of non-photosynthetic bradyrhizobia, can also nodulate *Aeschynomene indica* through their T3SS and independently of NFs [5]. This highlights the interest of *A. indica* to study these two alternative symbiotic processes.

The T3SS is a complex secretory machinery resembling a nanosyringe that is used by various Gram-negative bacteria to inject into the host cells type III effector proteins (T3Es) whose most common functions are to subvert host immunity [7, 8]. Rhizobial T3Es often act similarly as pathogenic T3Es and, depending on the host plant, can display opposing roles [9–11]. In some plants, they can repress the plant immune system and

[1]IRD, Laboratoire des Symbioses Tropicales et Méditerranéennes (LSTM), UMR IRD/Institut Agro/INRAE/Université de Montpellier/CIRAD, TA-A82/J– Campus de Baillarguet 34398, Montpellier cedex 5, France. [2]PHIM Plant Health Institute, Université de Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France. [3]University of Cambridge, Sainsbury Laboratory (SLCU), Cambridge CB2 1LR, UK. [4]LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France. [5]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France. [6]Graduate School of Agriculture, Tokyo University of Agriculture and Technology, Tokyo 183-8509, Japan. [7]Carrefour International, Bureau Régional Afrique de l'Ouest, Dakar, Sénégal. [8]Embrapa Agrobiologia, Bairro Ecologia, Seropedica, Rio de Janeiro, Brazil. [9]School of Life Sciences and State Key Laboratory of Agrobiotechnology, The Chinese University of Hong Kong, Hong Kong SAR, China. ✉email: eric.giraud@ird.fr

Received: 24 April 2023 Revised: 8 June 2023 Accepted: 13 June 2023
Published online: 24 June 2023

thus promote the infection and nodulation, whereas in other plants, they can be recognized by cytoplasmic plant receptors and trigger a strong plant immune response, called effector-triggered immunity (ETI), resulting in blockage of the rhizobial infection. Specific rhizobial T3Es can also circumvent the requirement of the NF signal and directly activate nodule formation. This is the case for the ErnA effector identified in *Bradyrhizobium vignae* strain ORS3257 [12]. This T3E, which is found in a wide range of *Bradyrhizobium* strains having a T3SS, plays a crucial role during the T3SS-dependent symbiosis with *Aeschynomene* [12]. Furthermore, ectopic expression of *ernA* in transgenic roots of *A. indica* triggers nodule-like structures in the absence of the bacteria [12]. The ErnA protein displays no homology with known functional domains, but it has been shown that ErnA targets the host nucleus and interacts with nucleic acids suggesting that ErnA could transcriptionally activate the nodulation program [12].

In the case of the T3SS-dependent symbiosis between *G. max* and *B. elkanii* USDA61, it is another T3E, named Bel2-5, which governs the interaction [13, 14]. Bel2-5 is also targeted to the host cell nucleus and contains a C-terminal small ubiquitin-like modifier (SUMO)-Protease domain shown to be required for the nodulation activity. This indicates that this T3E fulfills its role by deSUMOylation of plant protein targets. Taken together, these data indicate that different classes of rhizobial T3Es able to trigger nodulation and proposed to be called ET-Nods [15], use different mechanisms to activate the nodulation program in legumes.

In this study, we examined the ability of a collection of 196 *Bradyrhizobium* strains of various origins to use a NF-independent symbiotic process for nodulation of *A. indica*. We focused on genome-sequenced strains to establish correlations between genomic content and symbiotic properties. This led us to examine the following questions: are there phylogenetic groups distinct to photosynthetic bradyrhizobia able to develop a NF-independent, T3SS-independent symbiosis? Does the presence of ErnA in T3SS containing strains systematically confer the ability to develop a T3SS-dependent symbiosis? Can other T3Es distinct from ErnA act similarly and trigger nodulation in *A. indica*? Through a combination of comparative genomic analyses and mutagenesis of some candidate genes emerging from these analyses, we identified several novel ET-Nods and obtained more information about the ability of bradyrhizobia to nodulate legumes independently of NFs.

MATERIALS AND METHODS

Bacterial strains and growth conditions

All tested bacterial strains are listed in Supplementary Dataset S1. The *Bradyrhizobium* strains were grown in YM or AG medium at 28 °C [16, 17]. All strains were checked for purity by streaking on agar plate and for identity by ITS sequencing as described [18]. *Escherichia coli* strains were grown at 37 °C in modified Lysogeny Broth (LB) medium [19]. *Agrobacterium rhizogenes* Arqua1 and *Agrobacterium tumefaciens* GV3101 were grown at 28 °C on AG medium and *Ralstonia pseudosolanacearum* was grown in rich BG medium [20]. When required, the media were supplemented with the appropriate antibiotics at the following concentrations: 50–100 µg/mL kanamycin, 20 µg/mL nalidixic acid, 20 µg/mL cefotaxime and 100–200 µg/mL spectinomycin.

Genome sequencing and phylogenomic analysis

The genomes of strains ORS86, ORS111, STM3557, STM3561, STM3562, STM3566, LMG29515 and LMG09283 were sequenced by combining long read and short read sequencing technologies as previously described [21], assembled using the CulebrONT pipeline [22] and annotated using the MicroScope platform [23]. The other genomes downloaded from the NCBI or GenBank databases were annotated de novo by the MicroScope pipeline, and are also publicly available from the MicroScope platform (http://www.genoscope.cns.fr/agc/microscope). From all these genome sequences, a phylogenomic tree was constructed using IQ-tree v2.2.0.3 [24] based on a multi sequence alignment of 725 genes of the core genome shared by at least 98% of *Bradyrhizobium* strains, generated by PPanGGolin v1.2.63 [25]. The phylogeny was inferred from the resulting alignment under the model recommended by ModelFinder [26] and branch supports with ultrafast bootstrap were estimated from 100,000 iterations [27]. Final editing was carried out using iTOL (https://itol.embl.de/) [28].

Prediction of T3E repertoires, search of SUMO-proteases and phylogenetic analysis

The search of T3Es was performed as previously described [12]. The search for SUMO-Proteases was performed by interrogating the precomputed InterProScan annotations hosted by the MicroScope platform for each genome, using the InterPro identifier (IPR003653) corresponding to the C-terminal catalytic domain of the ULP1 protease family. Protein alignments of ErnA and SUMO-Proteases were performed using ClustalW (http://www.clustal.org/omega/) [29] and phylogenetic analysis using IQ-tree as described above.

Plasmid construction, mutagenesis and complementation

Standard molecular biology techniques were used for all cloning work. All constructions made in this study are listed in Supplementary Dataset S2, which also includes the primers and the cloning strategies. Briefly, the insertional mutants were obtained by single-crossing over. For this purpose, a 200 to 300-bp internal fragment of the target gene was amplified by polymerase chain reaction (PCR) and cloned into a derivative of the non-replicative plasmid pVO155 carrying an antibiotic resistance marker appropriate to the strain to be mutated (pVO155-npt2-GFP-npt2-Cefo for NAS96.2 and LMTR13 strains; pVO155-npt2-GFP-npt2-Sp for WSM1744 and ORS86) [5, 30]. The deletion WSM1744Δ*sup3* mutant was obtained by double crossing over. For this, 750 to 1000-bp PCR fragments corresponding to the upstream and downstream flanking regions of the $sup3_{WSM1744}$ gene (JAAVLW01_10281) were merged by overlap extension PCR and cloned into the pNPTS129 vector, which carries the *sacB* suicide gene [31]. Subsequently, a spectinomycin resistance cassette was introduced between the upstream and downstream flanking regions. The resulting plasmid was then transferred by conjugation into WSM1744 and double recombinants were selected as described previously [12].

For complementation of the WSM1744Δ*sup3* mutant, a PCR-fragment encompassing the full $sup3_{WSM1744}$ gene and the 500-bp upstream promoter region obtained with the high-fidelity enzyme Phusion DNA polymerase (Thermo Fischer Scientific) was cloned into the pVO155 vector [32] and re-introduced into WSM1744Δ*sup3* by single-crossing over.

For introduction of $ernA_{ORS3257}$ into USDA110 and USDA61 strains, PCR fragments corresponding to the promoter regions of $ernA_{USDA110}$ or $ernA_{USDA61}$ and the full $ernA_{ORS3257}$ gene were merged by overlap extension PCR, subsequently cloned into the pVO155 vector [32], and introduced into USDA110 or USDA61 strains by single-crossing over.

For complementation experiments using the ORS3257Δ*ernA* mutant, full-length $ernA_{USDA110}$, $ernA_{USDA61}$, $sup3_{WSM1744}$, $sup3_{NAS96\text{-}2}$ and $bel2\text{-}5_{USDA61}$ were amplified as described above and cloned downstream of the $ernA_{ORS3257}$ promoter into pVO155-pm-*ernA* (Supplementary Dataset S2) before re-introduction into ORS3257Δ*ernA* by single-crossing over.

For site-directed mutagenesis of *ernA*, we used the plasmid pVO155-pm-*ernA*::$ernA_{ORS3257}$ previously constructed to complement the ORS3257Δ*ernA* mutant [12], the Q5 Site-Directed Mutagenesis Kit (NEB) and the primers indicated in Supplementary Dataset S2.

Plant cultivation and symbiotic analysis

A. indica plants were cultured as previously described [5]. Eight plants per strain were inoculated and the symbiotic properties of the strain (number of nodules per plant, nitrogenase activity, infection of the nodules) were analyzed 21 days after inoculation as previously described [33]. The experiments were carried out at least in duplicate.

Sup3 secretion analysis

$Sup3_{WSM1744}$, *GFP* and *PopP2* coding sequences were introduced by LR recombination in pLAFR6-P2GFH/*hrpB* vector for expression of the protein of interest fused with a C-terminal 3x-HA epitope tag under the control of the HrpB-activated popP2 promoter sequence (Supplementary Dataset S2). The derived recombinant plasmids were introduced into *Ralstonia pseudosolanacearum* Δ*popP2* strain (GRS100 [34]) and Δ*popP2*/ Δ*hrcv* [35] strains by electroporation and selected on 5 µg/mL tetracycline and

7.5 µg/mL gentamicin. Production of *R. pseudosolanacearum* concentrated supernatants (from 10 mL of culture) was performed as previously described [36] with minor modifications. Briefly, transformed bacteria were grown in MP secretion medium (MP medium 1×, Congo Red 100 µg/mL, glutamate 20 mM and tetracycline 5 µg/mL) for 30 h at 28 °C under shaking. Pelleted cells from 1 mL of culture were lysed in 300 µL of 2× Laemmli buffer (pellet fraction). Proteins from the cell culture supernatant were precipitated with TCA 25% (v/v) for 12 h at 4 °C and resuspended in 50 µL of 2× Laemmli buffer. Protein extracts were denaturated for 5 min at 95 °C, loaded on 10% SDS-PAGE gels, and transferred on nitrocellulose membrane using Trans-Blot Turbo transfer kit (Bio-Rad) according to the manufacturer's instructions. Blotted membranes were blocked with 5% milk in Tris-buffered saline plus 0.01% Tween 20 (TBS-T) for one hour at room temperature and subsequently incubated with desired antibodies at 4 °C overnight. Antibodies used were anti-HA-HRP (3F10, Roche, dilution 1:5000), polyclonal rabbit anti-GroEL (Stressgen Biotechnologies Corporation, dilution 1:20,000) and goat anti rabbit IgG-HRP (Bio-Rad, dilution 1:10,000). Chemiluminescent horseradish peroxidase (HRP) detection was performed using the Clarity Western ECL substrate (Bio-Rad).

Agrobacterium tumefaciens infiltration assays in *Nicotiana benthamiana*

Binary plasmid *pB7FWG2-35S-eGFP-3Flag-sup3*$_{WSM1744}$ used for constitutive expression of eGFP-3Flag-Sup3 (Supplementary Dataset S2), was introduced into *A. tumefaciens* strain GV3101 by electroporation [37]. Agro-infiltration of *N. benthamiana* leaves and localization of fluorescently labeled Sup3 were performed as previously described [12].

Hairy root transformation with *Agrobacterium rhizogenes*

Plasmid *p35S-sup3* containing *sup3*$_{NAS96\text{-}2}$ under the control of the 35S promoter (Supplementary Dataset S2), and the empty vector pJCV51 with the DsRed marker (https://gateway.psb.ugent.be), were introduced by electroporation into the *A. rhizogenes* Arqua1 strain used for hairy root transformations. *A. indica* root transformation was performed following previously described procedures [38].

RESULTS

A large diversity of *Bradyrhizobium* strains nodulate *A. indica*

We inoculated *A. indica* with 196 *Bradyrhizobium* strains and found that 69 strains induced nodules or nodule-like structures (Supplementary Dataset S1). Six distinct symbiotic phenotypes (named Type A to F) could be distinguished (Fig. 1). The Type A was the largest with 127 strains unable to induce nodules (Nod$^-$). The 30 Type B strains showed a very low nodulation capacity, only a part of the inoculated plants displayed a low number (<5 per plant) of nodule-like structures and infection of these nodules was rare and limited to the intercellular spaces. The 16 Type C strains, including USDA61, induced a large number of nodules in all inoculated plants but the infection remained limited to the intercellular spaces as in Type B. Type D was observed for only three strains (BR2003, ORS86 and ORS111), and is characterized by a multilobed unorganized nodule development leading to numerous protuberances all along the root, which look more like

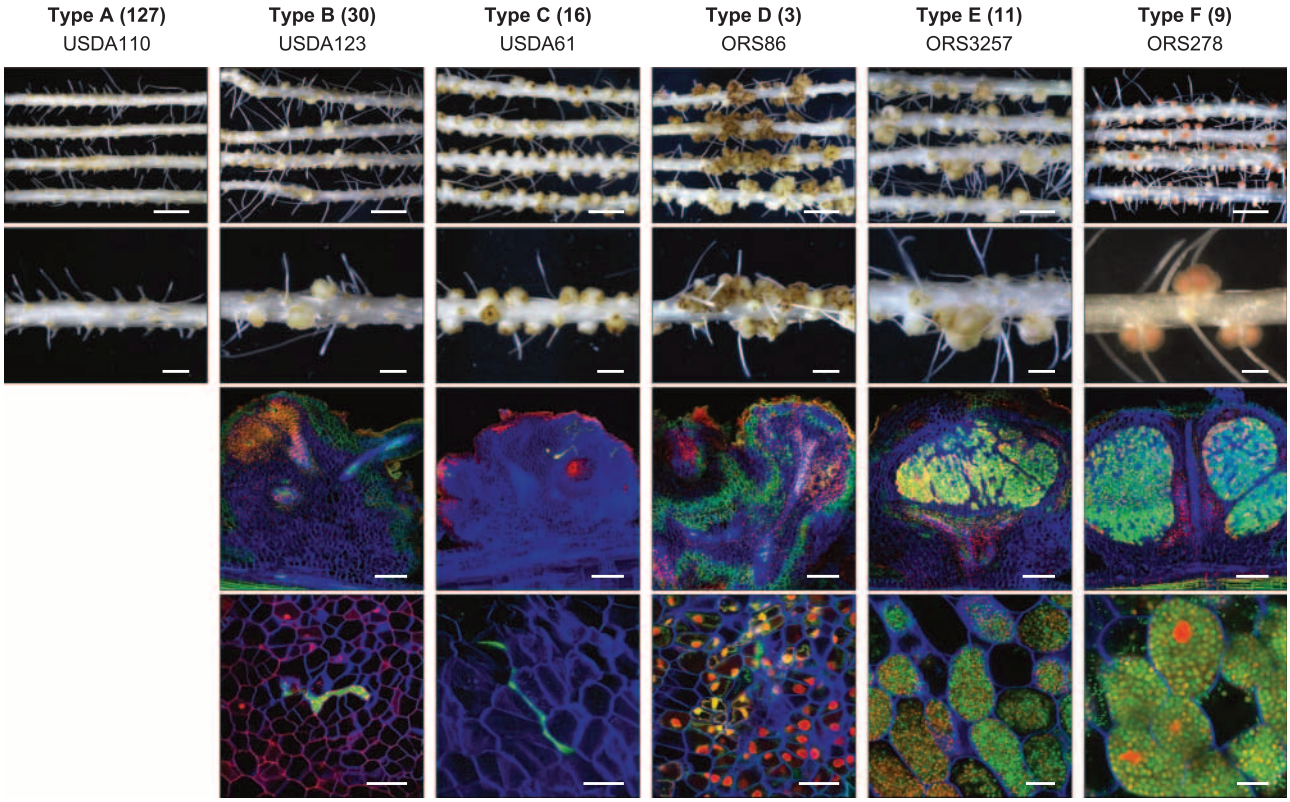


Fig. 1 Symbiotic phenotypes of various *Bradyrhizobium* strains inoculated on *A. indica*. Six different symbiotic phenotypes were distinguished and designated as Type A to F (from the least to the most efficient symbiotic interaction) and for each phenotype a representative strain is shown. Type A: strains unable to nodulate. Type B: strains showing a weak nodulation (average < 5 nodules per plant) with sometimes intercellular infection. Type C: strains inducing a large number of nodules (average > 20 nodules per plant) with intercellular infection. Type D: strains forming massively multilobed nodules with intercellular infection. Type E: strains inducing misshapen and large nodules with a zone of intracellular infected plant cells but having a weak nitrogenase activity (no bacteroid differentiation). Type F: round-shaped nodules with a central zone containing intracellular infected plant cells with bacteria differentiated into bacteroids that efficiently fix nitrogen for the benefit of the plant. The number of strains in each category is shown in brackets. Row 1 and 2: photo of roots and nodules induced by the different strains. Scale bars: row 1, 0.5 cm; row 2, 0.2 cm. Row 3 and 4, Confocal microscopy images of section of nodules induced by the different strains observed after staining with SYTO 9 (green, live bacteria), propidium iodide (red, infected plant nuclei and dead bacteria or bacteria with compromised membranes) and calcofluor (blue, plant cell wall); Scale bars: row 3, 200 µm; row 4, 40 µm except for ORS3257 and ORS278 (10 µm).

galls than nodules. The presence of intercellular bacteria could sometimes be detected in these structures. In Type E (11 strains), which includes ORS3257, the elicited nodules were large, generally misshapen but with a clear intracellularly infected central tissue. However, as previously described [5], the intracellular bacteria did not differentiate properly into bacteroids. This is correlated with no or very low nitrogenase activity (Supplementary Dataset S1), which in all cases was insufficient to support plant growth (Supplementary Fig. S1). Finally, the Type F (9 strains), mainly represented by photosynthetic *Bradyrhizobium* strains, corresponds to bacteria inducing regular round-shaped nodules, infected intracellularly by bacteria differentiated into spherical bacteroids [39], which efficiently fixed nitrogen for the plant benefit. Thus, the ability to initiate nodule organogenesis in *A. indica* is widespread throughout the *Bradyrhizobium* genus but the outcome of this interaction strongly varies among the strains.

Phylogenetic distribution of *nod*, *T3SS* and *ernA* genes in relation with the symbiotic properties

The assayed *Bradyrhizobium* strains cover the seven phylogenetic clades previously identified within the *Bradyrhizobium* genus and named *B. japonicum*, Photosynthetic, Kakadu, *B. elkanii*, *B. jicamae*, Soil 1 and Soil 2 Supergroups (Fig. 2) [40]. Correlating symbiotic properties with phylogenetic position showed that the Type F phenotype was restricted to the photosynthetic supergroup. Most of the strains of this group lack *nod* and *T3SS* genes. All other bacteria (59/69 strains) nodulating *A. indica* are scattered in 4 of the 6 other supergroups and all possess a classical T3SS of the Rhc-I type, except for the strain Tv2a-2 which has only an atypical T3SS [41]. Taken together, these observations suggest that the NF-independent, T3SS-independent symbiotic process is specific to strains belonging to the photosynthetic supergroup whereas the ability to use a T3SS-dependent process to induce nodulation is a more widespread trait found in a large variety of *Bradyrhizobium* strains.

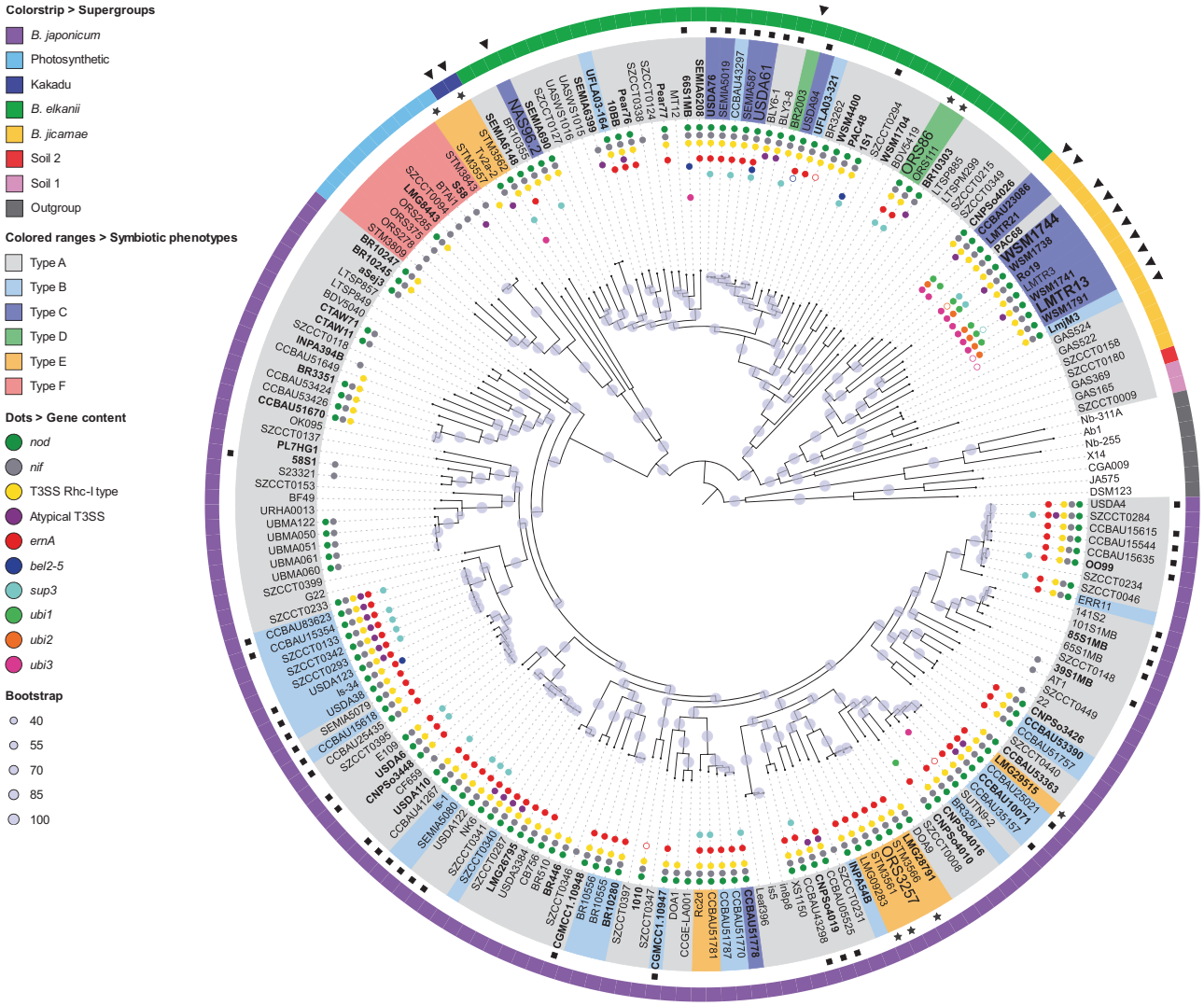


Fig. 2 Maximum-likelihood phylogenomic tree of *Bradyrhizobium* strains tested in this study. The phylogeny was inferred from the core genome of the 196 *Bradyrhizobium* strains tested and several species from *Rhodopseudomonas* and *Nitrobacter* used as an outgroup. Purple circles on the nodes indicate ultrafast bootstrap values calculated by IQ-Tree. The presence of some symbiotic determinants in each strain is shown by colored dots referenced in the key, full dots indicate complete genes, empty dots indicate partial genes. Type strains are indicated in bold. All previously described supergroups [40] are present in our analyzed collection and indicated by a band strip color referenced in the key: *B. japonicum*, photosynthetic, Kakadu, *B. elkanii*, *B. jicamae*, Soil 1 and Soil 2. The symbiotic phenotype (Type A to F) of each strain is indicated by a colored range referenced in the key. The black triangles indicate strains inducing nodules while they lack *ernA*. The black stars indicate strains for which the genome has been sequenced in the frame of this study. The black squares indicate strains isolated from soybean nodule. The name of strains with a higher police size are those used for functional analyses in this study.

We found that 45 of the 59 nodulating strains have an ErnA homolog (Supplementary Dataset S3), supporting the idea that ErnA is a widespread T3E in the *Bradyrhizobium* genus to trigger nodulation. However, correlation between the presence of ErnA and the ability to nodulate is not strict, as 33 strains having a complete ErnA homolog were Nod$^-$ and on the contrary, 14 strains lacking an ErnA homolog nodulated *A. indica* (Fig. 2).

Effect of ErnA variability on nodulation and infection

ErnA variants exist that differ by the presence or not of a 80 AA domain at its N-terminal end [12]. Furthermore, AA variations are

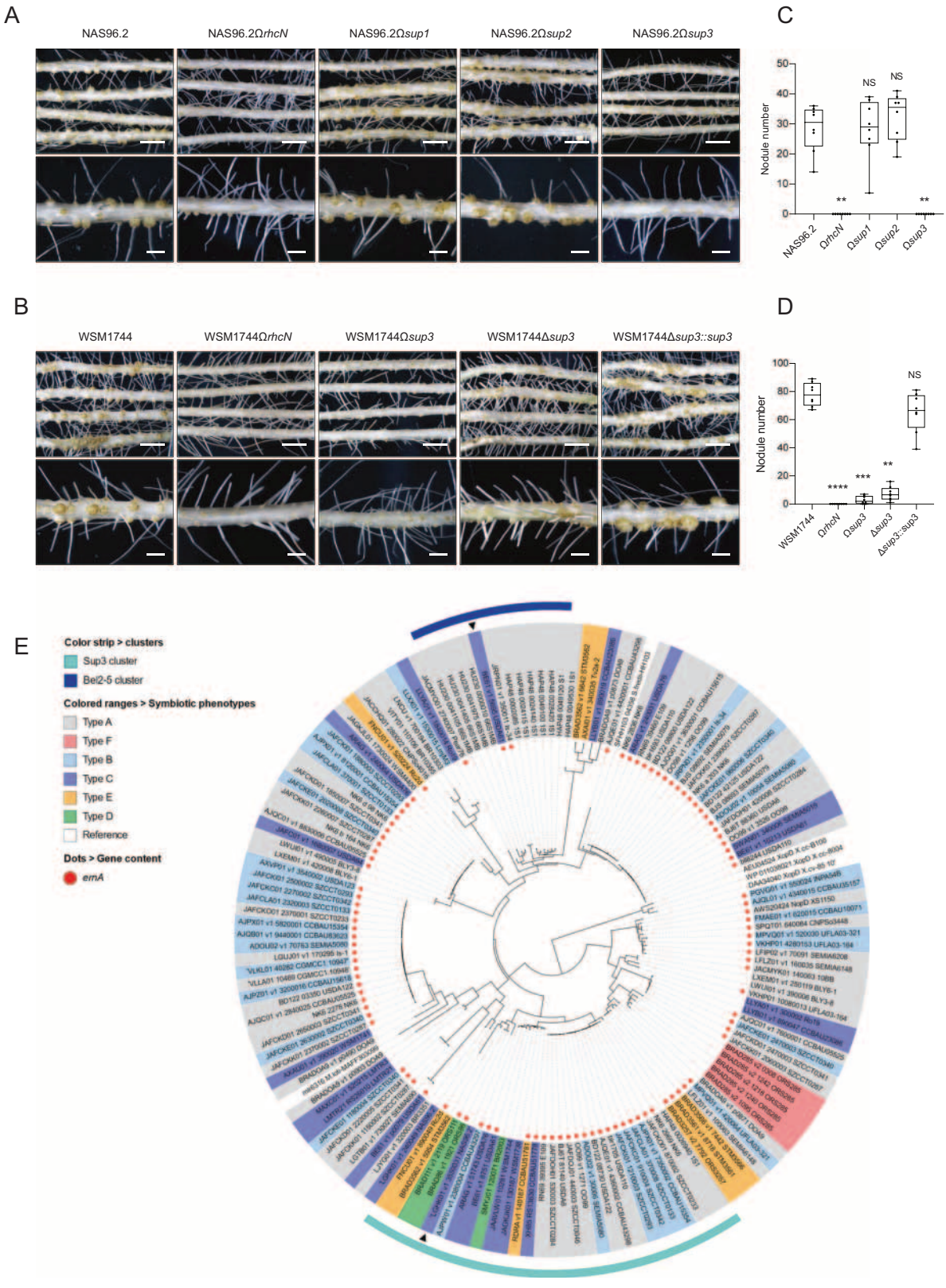


Fig. 3 Sup3, a putative SUMO-protease conferring the ability to nodulate *A. indica*. Images of roots and nodules of *A. indica* plants at 21 days post inoculation with strain NAS96.2 and its mutant derivatives in T3SS (Ω*rhcN*) and in different genes encoding putative SUMO-proteases (Ω*sup1*, Ω*sup2* and Ω*sup3*) (**A**) and strain WSM1744 containing a *sup3* homolog and its mutant derivatives (Ω*rhcN*, Ω*sup3*, Δ*sup3* and Δ*sup3::sup3)* (**B**) (scale bars: upper panels 0.5 cm, lower panels 0.2 cm). Nodule number of *A. indica* plants at 21 days after inoculation with NAS96.2 (**C**) or WSM1744 (**D**) and their respective mutants. Box plots show the results of one of the two experiments performed independently (8 plants each). $^{**}p \leq 0.001$, significant differences between the wild-type strain NAS96.2 and each mutant (**C**) and $^{**}p \leq 0.005$, $^{***}p < 0.0005$, $^{****}p < 0.0001$, significant differences between the wild-type strain WSM1744 and each mutant (**D**) using a nonparametric Kruskal–Wallis test, NS not significant. **E** Phylogenetic tree of full-length SUMO-proteases identified in the collection of *Bradyrhizobium* strains used in this study. Considering the very different branch lengths, the MAD (minimal ancestor deviation) method was used for rooting the tree [59]. The SUMO-proteases Sup3 from NAS96.2 and Bel2-5 from USDA61 (indicated by black triangles) belong to two distinct phylogroups (indicated by a color strip). The symbiotic phenotype of the strains in which the SUMO-protease are identified is indicated by a colored range referenced in the key and the presence of an *ernA* homolog by a red circle. As references (white range), the following SUMO-proteases are included: (i) mlr6316 identified in *Mesorhizobium loti* MAFF303099 [60]; (ii) SFHH103_04358 identified in *Sinorhizobium fredii* HH103 [61] and (iii) The XopD SUMO-proteases [57] identified in various *Xanthomonas campestris* strains.

also observed between ErnA orthologs. A phylogenetic tree of ErnA highlights its diversity (Supplementary Fig. S2A). In the branch that groups the highest number of ErnA proteins (34), a very high level of AA conservation is observed. Most of the strains of this phylogroup failed to nodulate (21 strains out of 34), while the remaining 13 showed a very weak nodulation ability (Type B phenotype), raising the question whether this form of ErnA is active. To analyze this, we introduced the *ernA* gene from the USDA110 strain ($ernA_{USDA110}$) into the ORS3257Δ*ernA* mutant which lost its nodulation ability [12]. The addition of the original *ernA* gene of ORS3257 restored nodulation while the introduction of $ernA_{USDA110}$, put under the control of the $ernA_{ORS3257}$ promoter, resulted only in weak nodulation (Supplementary Fig. S2B). This indicates that $ErnA_{USDA110}$ is less active than $ErnA_{ORS3257}$. Furthermore, when the *ernA* gene from ORS3257 was introduced into USDA110, the transformed strain (USDA110::$ernA_{ORS3257}$) did not acquire nodulation ability (Supplementary Fig. S2D), indicating that factors other than ErnA are responsible for the Nod^- character of USDA110.

Important differences in the nodule infection were also observed between strains using a T3SS-dependent process, with some inducing only intercellularly infected nodules (Type B, C or D) and others inducing intracellularly infected nodules (Type E). We investigated whether these differences could be related to variations in the ErnA sequence, and thus whether ErnA could control both nodulation and infection as NFs do. To address this question, we introduced the *ernA* gene from USDA61 ($ernA_{USDA61}$) (Type C) into the Δ*ernA* mutant of ORS3257 (Type E). No difference in nodule number and the type of infection was observed between plants inoculated with WT ORS3257 and the ORS3257Δ*ernA*::$ernA_{USDA61}$ mutant (Supplementary Fig. S2B, C). Similarly, introduction of $ernA_{ORS3257}$ into USDA61 did not promote intracellular infection (Supplementary Fig. S2E). This data suggests that the main function of ErnA is dedicated to nodule organogenesis but not infection.

Strain NAS96.2 contains a novel putative ET-Nod

As mentioned above, 14 strains lacking an *ernA* homolog can nodulate *A. indica* (Fig. 2). This raises the question: Do these strains use a T3SS-independent mechanism or an alternative T3E that mimics the action of ErnA? We focused on the NAS96.2 strain, which is closely related to three strains (SEMIA6148, SEMIA690, BR1035) unable to nodulate *A. indica* but having a T3SS (Fig. 2), suggesting the possibility that NAS96.2 has a specific ET-Nod that is absent in these three related strains. By comparing their putative effectome (Supplementary Dataset S4), we identified three strain-specific T3Es in NAS96.2, that we annotated Sup1, Sup2 and Sup3, considering that all of them encode putative SUMO-proteases. Insertional mutants of these three genes were constructed as well as a T3SS mutant (insertion mutant in the *rhcN* gene). The *rhcN* mutation abolished nodulation, confirming that the process is dependent on T3SS (Fig. 3A, C). Furthermore, while the *sup1* and *sup2* mutants behave like the WT strain, the *sup3* mutant did not induce nodules (Fig. 3A, C), suggesting that in the NAS96.2 strain, Sup3 is an ET-Nod controlling nodulation of *A. indica*.

A sub-class of SUMO-proteases in *Bradyrhizobium* genus confers nodulation ability

The *sup3* gene (LGHK01_v1_3550031) encodes a very large protein of 1463 AA which displays some similarities with Bel2-5 and carries at its C-terminus, a SUMO-protease domain of the C48 peptidase [ubiquitin-like protease 1 (ULP1)] family. Considering that SUMO-proteases were previously identified as one of the largest classes of T3Es in bradyrhizobia [41] and that two of them (Bel2-5 and Sup3) could act as ET-Nod, we examined in more details the distribution of SUMO-proteases in our strain collection. By searching for the presence of the ULP1 domain in the proteomes of these bacteria, we identified a total of 228 SUMO-proteases present in 91 strains, all of them containing a T3SS reinforcing the idea that these SUMO-proteases correspond to T3Es (Supplementary Dataset S5).

A phylogenetic analysis on the 149 available full-length SUMO-protease sequences (Supplementary Dataset S5) revealed that Bel2-5 and Sup3 belong to two distinct clades (Fig. 3E). Plotting on this tree the symbiotic properties of the corresponding strains reveals that strains of the Sup3 cluster predominantly nodulated *A. indica*. Furthermore, among these strains, four of them, in addition to NAS96.2, lack *ernA*, suggesting that the presence of a Sup3-like SUMO-protease might confer the ability to nodulate *A. indica*. To support this hypothesis, we constructed *sup3* (Ω*sup3)* and T3SS (Ω*rhcN)* insertional mutants in the *B. archetypum* WSM1744 strain. The T3SS mutation completely aborted the nodulation and the *sup3* mutation also resulted in a drastic reduction of nodules (Fig. 3B, D). Furthermore, the Δ*sup3* deletion mutant had the same phenotype as the Ω*sup3* insertional mutant and re-introduction of $sup3_{WSM1744}$ restored nodulation to WSM1744 wild-type level (Fig. 3B, D) confirming that the nodulation deficiency of *sup3* mutants was specifically due to *sup3* inactivation. Taken together, these data indicate that the nodulation between WSM1744 and *A. indica* is T3SS-dependent and that Sup3, as in NAS96.2, is the main T3E governing nodule organogenesis.

Several *Bradyrhizobium* strains contain multiple putative ET-Nods

Among the group of strains possessing a *sup3* homolog, we observed that some have additional putative ET-Nods. This is the case of ORS86 which carries a *sup3* and *ernA* homolog, and USDA61 which contains in addition to *sup3* and *ernA* also *bel2-5* (Fig. 2). This raises the question whether these different ET-Nods have cumulative effect in nodule induction, especially in the case of ORS86 which induced abundant nodulation in *A. indica* (Fig. 1). Analysis of the symbiotic properties of the ORS86 and USDA61 mutants in these different ET-Nods and in the *rhcN* gene (Fig. 4),

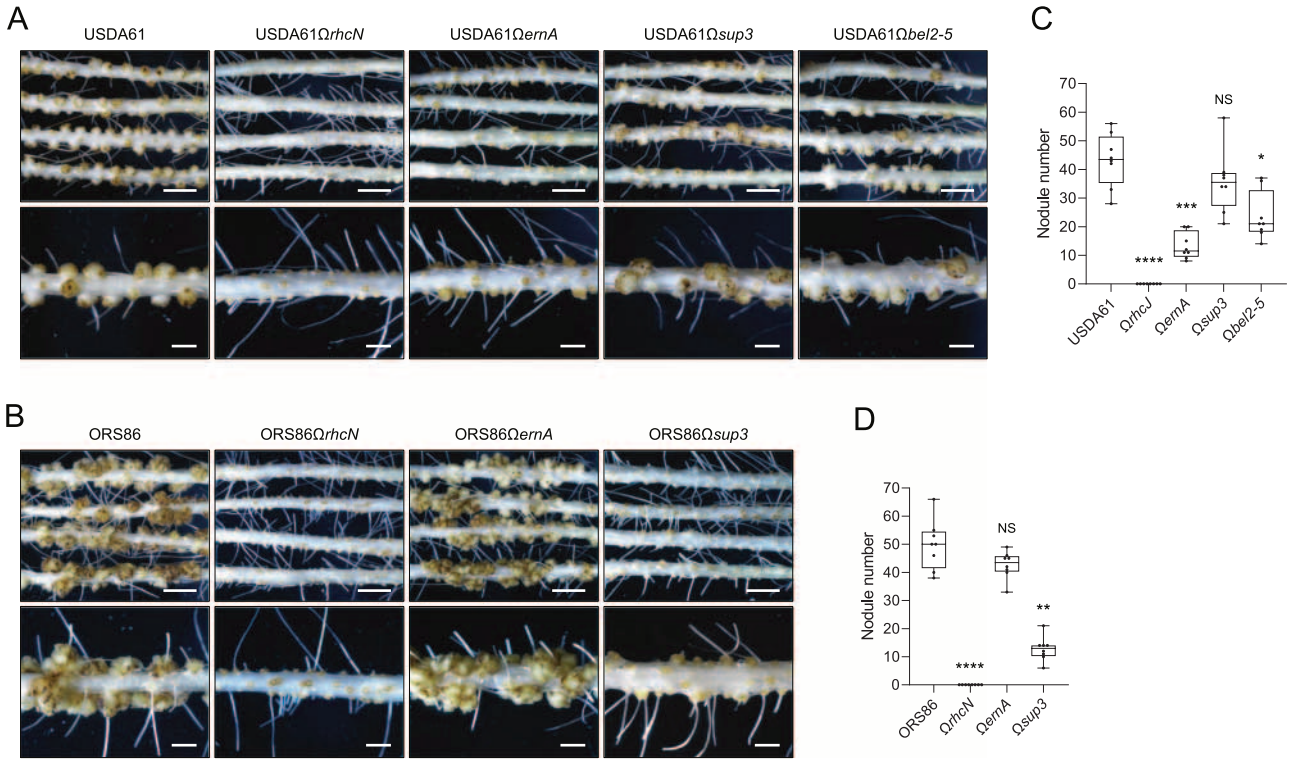


Fig. 4 The ET-Nod governing nodulation differs according to the strain. Images of roots and nodules of *A. indica* plants at 21 days post inoculation with strain USDA61 (**A**) or ORS86 (**B**) and their respective mutants in T3SS (Ω*rhcJ* or Ω*rhcN*) and several identified ET-Nods (Ω*ernA*, Ω*sup3*, Ω*bel2-5*) (scale bars: upper 0.5 cm, lower 0.2 cm). Nodule number on *A. indica* plants at 21 days after inoculation with USDA61 (**C**) or ORS86 (**D**) and their respective mutants. Box plots show the results of one of the two experiments performed independently (8 plants each). $*p < 0.05$, $**p < 0.005$, $***p \leq 0.0005$, $****p < 0.0001$, significant differences between the wild-type strain and each mutant using a nonparametric Kruskal–Wallis test, NS not significant. The mutants USDA61Ω*ernA*, USDA61Ω*rhcJ* and USDA61Ω*bel2-5* were previously obtained [12, 62, 63].

showed that i) the abundant nodulation of ORS86 was T3SS-dependent, as observed in USDA61 (Fig. 4A, B), and ii) only one of the putative ET-Nods plays a prominent role in nodulation, Sup3 in the ORS86 strain and ErnA in USDA61 (Fig. 4C, D).

During the interaction between USDA61 and *Glycine max* cv Enrei, it was shown that Bel2-5 controls nodulation, whereas the ErnA homolog plays only a minor role [13]. Thus, the ET-Nod governing the nodulation process is strain-dependent but also host plant-dependent.

Introduction of *sup3*, but not *bel2-5*, restores nodulation by ORS3257Δ*ernA*

To further demonstrate that Sup3 acts as an ET-Nod, we used a gain of function approach by analyzing the ability of Sup3 to complement the nodulation-deficient ORS3257Δ*ernA* mutant. For this purpose, the *sup3* gene from WSM1744 and NAS96.2 under the control of the *ernA* promoter was introduced into ORS3257Δ*ernA*. We also tested the complementation with the *bel2-5* gene from USDA61. Introduction of $sup3_{WSM1744}$ or $sup3_{NAS96.2}$ but not of *bel2-5* restored the ability of the mutant to induce intracellularly infected nodules on *A. indica* (Fig. 5). In both cases the number of formed nodules was greater than obtained with WT ORS3257 (Fig. 5A, B). This data confirms the ability of Sup3 to trigger nodulation on *A. indica*, which in the ORS3257 background, seems to be higher than ErnA.

The ectopic expression of *sup3* in *Aeschynomene* spp. roots induces spontaneous nodules

To analyze if Sup3 alone can induce nodule organogenesis, transgenic lines overexpressing $sup3_{NAS96.2}$ were generated in two *Aeschynomene* species (*A. indica* and *Aeschynomene evenia)*. Seven weeks after *Agrobacterium rhizogenes* transformation, about 30% and 40% of the transformed roots of *A. indica* and *A. evenia*, respectively, contained nodule-like structures with an average of 5 to 10 nodules per nodulated plant (Fig. 6D–F and Supplementary Fig. S3). These spontaneous nodules were observed at the level of the emergence of lateral roots and contained peripheral vascular bundles, similar to nodules induced by the bacterium (Fig. 6G, H and Supplementary Fig. S3). These nodule-like structures were not observed on roots transformed with the empty vector (Fig. 6A–C and Supplementary Fig. S3). Thus, ectopic expression of *sup3* is sufficient to activate the nodule organogenesis program in *Aeschynomene*.

Sup3 is a bona fide T3E that targets the host nucleus

To check whether Sup3 can be secreted through the T3SS machinery, we introduced a C-ter His_6-tagged version of Sup3 into WSM1744 and its Δ*T3SS* mutant. However, despite several attempts, we were unable to detect Sup3-His_6 in both culture supernatant and cell pellets suggesting that synthesis of the tagged protein is too low to be detected. The extremely slow growth of the WSM1744 strain and the very large size of the Sup3 protein could be the reason for the difficulties in detection. Therefore, we used as an alternative approach a Type 3 secretion assay originally developed in the pathogenic bacterium *Ralstonia pseudosolanacearum*. Here, $Sup3_{WSM1744}$ was fused with a C-terminal 3x-HA epitope tag and expressed in a Δ*popP2* strain that has been previously used to demonstrate the T3SS-dependent secretion of the *R. pseudosolanacearum* T3E PopP2 [36]. GFP and PopP2 were used as negative and positive controls, respectively. As expected, $Sup3_{WSM1744}$ was detected in the pellet and the supernatant fractions, indicating that, like PopP2,

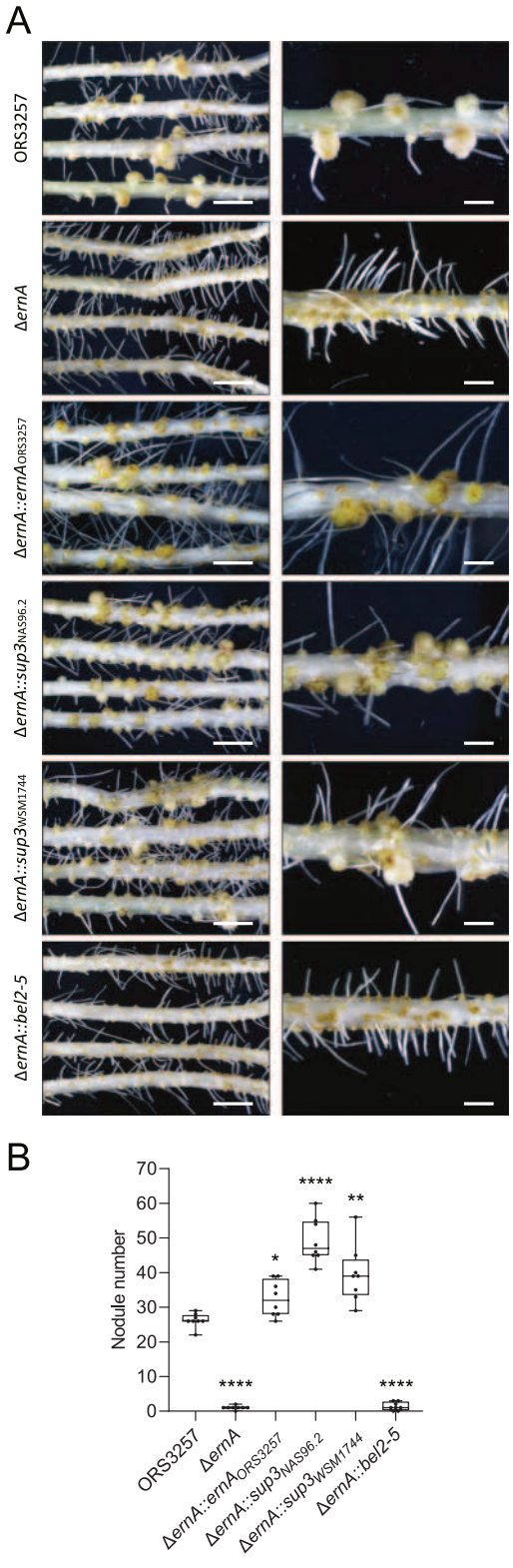


Fig. 5 *sup3*, but not *bel2-5*, complements nodulation of ORS3257Δ*ernA* mutant. A Symbiotic phenotypes of ORS3257 and various derivatives of the ORS3257Δ*ernA* mutant carrying *ernA*, *sup3* and *bel2-5* genes (Δ*ernA*::*ernA*$_{ORS3257}$; Δ*ernA*::*sup3*$_{NAS96.2}$; ORS3257Δ*ernA*::*sup3*$_{WSM1744}$; ORS3257Δ*ernA*::*bel2-5*) (scale bars: left 0.5 cm, right 0.2 cm). **B** Nodule number on *A. indica* plants at 21 days after inoculation with ORS3257 and the ORS3257Δ*ernA* mutant derivatives. Box plots show the results of one of the two experiments performed independently (8 plants each). *$p \leq 0.01$, **$p < 0.005$, ****$p < 0.0001$, significant differences between the wild-type strain and each mutant using a parametric Welch test.

Sup3$_{WSM1744}$ can be secreted through the T3SS (Fig. 7A). Of note, Sup3$_{WSM1744}$ and PopP2 were strongly processed after secretion by a mechanism that remains to be determined. By contrast, GFP that lacks a secretion signal was only detected in the pellet fraction. When expressed in a Δ*popP2*/Δ*hrcV* strain mutated both in *popP2* and in *hrcV* that encodes a conserved inner membrane component of the T3SS [42], Sup3$_{WSM1744}$ and PopP2 were absent in the culture supernatant confirming that their secretion was T3SS-dependent.

To investigate its subcellular localization *in planta*, Sup3$_{WSM1744}$ N-terminally fused with enhanced GFP and a triple hemagglutinin epitope tag (eGFP-3Flag-Sup3$_{WSM1744}$) was transiently expressed in tobacco leaf cells. As previously observed with ErnA [12], eGFP-3Flag-Sup3 was found to be targeted to the plant nucleus, suggesting that Sup3 triggers nodule organogenesis by interfering with host nuclear components (Fig. 7B).

Identification of three putative ET-Nods, Ubi1, Ubi2 and Ubi3, in *Bradyrhizobium icense* LMTR13

Except for the PAC68 strain, all ten tested strains from the *B. jicamae* supergroup nodulated *A. indica* while all lack an *ernA* homolog and most also lack a *sup3* homolog, suggesting the existence of other ET-Nod(s) within the strains of this phylogroup (Fig. 2). Comparative analysis of the effectome of these different strains identified only one common T3E (NopM) present specifically in nodulating strains (Supplementary Dataset S4). However, this T3E, which is among the most commonly found and studied T3Es in rhizobia [41], has never been reported to have a nodulation induction activity. To identify ET-Nods in these strains, we focused on the *B. icense* LMTR13 strain, which has the smallest effectome among the considered nodulating strains. We mutated the ten *T3E* genes, which are present in this strain but absent in PAC68 as well as the T3SS apparatus (*rhcN* mutant). Once again, we confirmed the critical role played by T3SS in this symbiotic process with the absence of nodulation observed for the *rhcN* mutant (Fig. 8A). Among the ten *T3E* mutants tested, 3 mutants in previously unknown *T3E* genes, which we named Ubi1, Ubi2 and Ubi3 (Ubi for Unknown T3Es identified in the *Bradyrhizobium icense* strain), gave a significant reduction in nodule number of 62, 40 and 42%, respectively, compared to the WT strain. The absence of a mutant with (nearly) completely abolished nodulation capacity, as observed for the *ernA* or *sup3* mutants in other strains, leads us to hypothesize that these three putative T3Es act synergistically.

Ubi1, Ubi2 and Ubi3 show no homology to known functional domains, but a nuclear localization signal can be predicted for each of them using Localizer [43] suggesting that these T3Es are also targeted to the host nucleus (Fig. 8B). Their distribution is relatively restricted; Ubi1, Ubi2 and Ubi3 homologs can be detected respectively in 8, 7 and 13 strains of the bradyrhizobia collection used here, among which a large proportion of strains belong to the *B. jicamae* supergroup (Fig. 2 and Supplementary Dataset S3).

Ubi1, Ubi2 and Ubi3 share common conserved motifs found in the C-terminus of ErnA

As the three new putative ET-Nods are proteins with unknown function, we used the MEME suite tools [44] to investigate if Ubi1, Ubi2 and Ubi3 share common motifs with other known and characterized T3Es. This analysis revealed that these three T3Es share two contiguous 21 AA motifs (M1 and M2), which are also found at the C-terminal end of ErnA (Fig. 8B and Supplementary Fig. S4). Furthermore, Ubi1 and ErnA share an additional 8 AA motif (M3) downstream of M1 and M2. This suggests a possible functional link between these different T3Es. The succession of M1 and M2, associated or not with M3, is also found in other T3Es (NopAR, InnB, NopL and NopAB) (Supplementary Fig. S4). The latter two were previously shown to play a role in the T3SS-dependent symbiosis [12, 13].

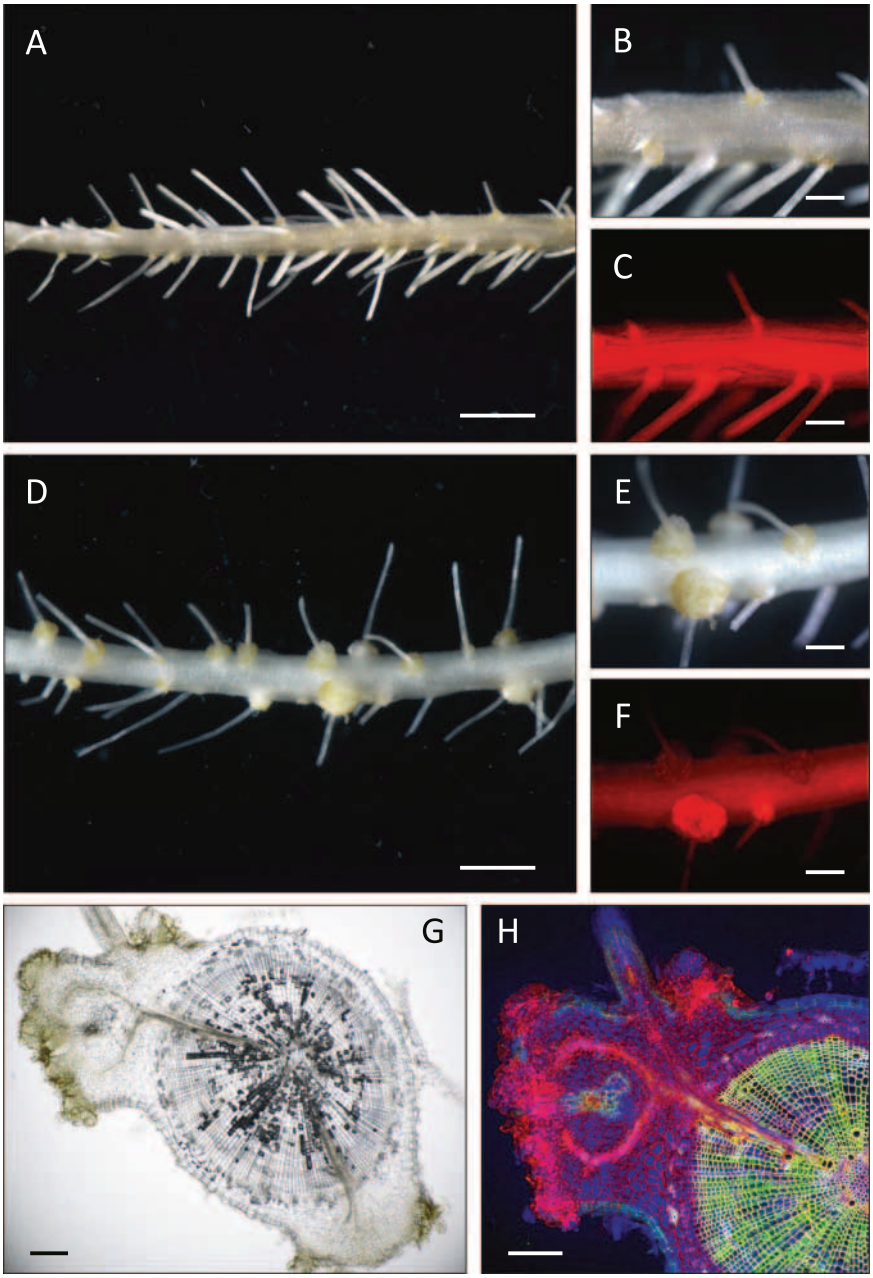


Fig. 6 Ectopic expression of *sup3* in *A. evenia* roots induces spontaneous nodules. *A. evenia* roots transformed either with the empty vector containing the DsRed marker (**A**–**C**) or with *p35S-sup3* (**D**–**F**) at 7 weeks after transformation (no bacterial inoculation). Roots were observed by a fluorescence stereomicroscope equipped with a DsRed filter. Scale bars: **A** and **D**, 2 mm; **B**, **C**, **E** and **F**, 500 µm. Micro-sections of pseudo-nodules observed using light microscopy showing the peripheral vascularization (**G**) and using a confocal microscope (**H**) after staining with SYTO 9, propidium iodide and calcofluor showing that the induced pseudo-nodules do not contain bacteria. Scale bars **G** and **H**, 200 µm.

The bioinformatic tool DP-bind [45] recognized the M2 motif as a DNA-binding site while the GPS-SUMO tool [46] predicted that the M3 domain is a SUMO interaction motif (SIM), a motif found in non-covalently SUMO-binding proteins [47]. Note that the M2 motif is found three times in ErnA (Fig. 8B). To gain further insight into the importance of these conserved motifs, we constructed different mutated forms of ErnA removing or modifying these motifs and tested their ability to complement the ORS3257Δ*ernA* mutant (Fig. 8C, D). We also included a mutagenesis of the SSS motif, which was also predicted as a putative DNA-binding site by DP-Bind. While ErnA$^{M2\text{-}1}$ and ErnA$^{SSS\text{->}AAA}$ restored nodulation ability of the ORS3257Δ*ernA* mutant, the ErnA variants mutated in the M1, M2-(2/3) and M3 motifs failed to do so (Fig. 8C, D). These data highlight the importance of these different motifs in ErnA functionality.

DISCUSSION

The NF-independent T3SS-independent symbiosis, a process restricted to the photosynthetic *Bradyrhizobium* supergroup

Using a large number of *Bradyrhizobium* strains, this study shows that the ability to nodulate *A. indica* via a NF-independent, T3SS-independent process is restricted to *Bradyrhizobium* strains belonging to the photosynthetic supergroup. The name of this group is misleading because it includes the non-photosynthetic strain STM3843, which also nodulates *A. indica* [48], showing

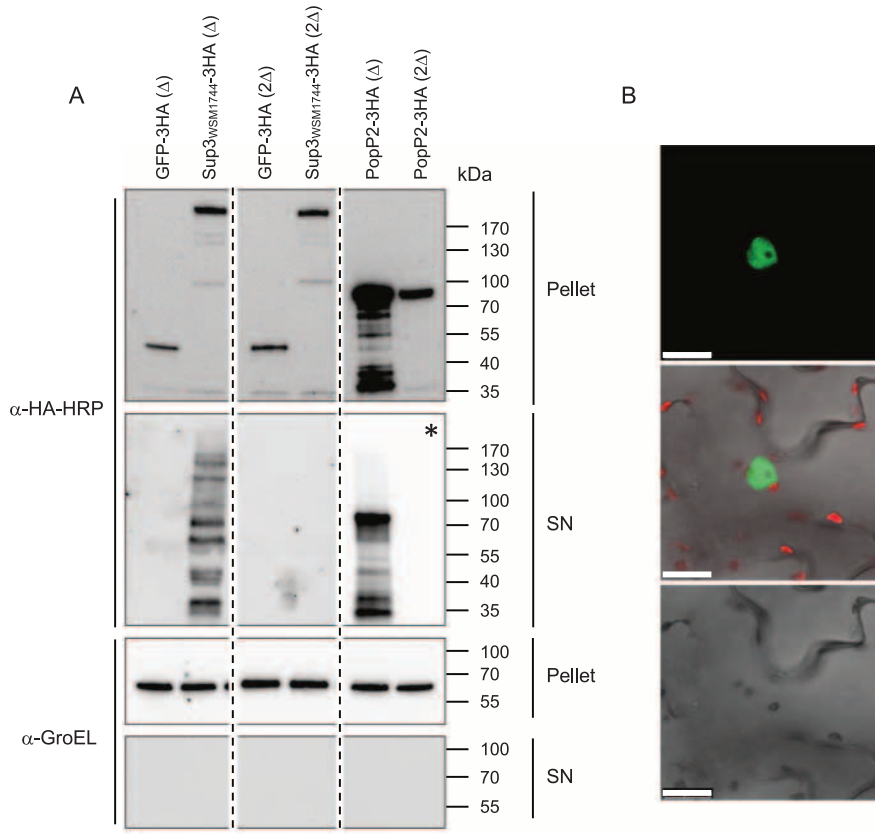


Fig. 7 Sup3$_{WSM1744}$ is secreted through the T3SS machinery of *Ralstonia pseudosolanacearum* and is targeted to the plant nucleus. A GFP-3HA, Sup3$_{WSM1744}$-3HA and PopP2-3HA either produced in Δ*popP2 (Δ)* or Δ*popP2*/Δ*hrcV (2Δ)* bacterial cells were immuno-detected from pellet and supernatant fractions using an anti-HA antibody. Like PopP2, Sup3$_{WSM1744}$ could not be detected in culture supernatants of *ΔpopP2/ΔhrcV* bacteria cells, demonstrating its proper secretion by the T3SS. All protein samples were probed with anti-GroEL antibody raised against the cytoplasmic chaperonin GroEL to ensure that no bacterial lysis had occurred during the preparation of supernatant fractions. The star (*) indicates a shorter exposure of the membrane with lanes corresponding to the supernatants of Δ*popP2* and Δ*popP2*Δ*hrcv* strains expressing PopP2-3HA. Dotted lines indicate unnecessary lanes that have been removed from the same membranes. **B** GFP fluorescence observed in *N. benthamiana* leaf cells transiently expressing eGFP-3Flag-Sup3$_{WSM1744}$. GFP fluorescence signal was visualized by confocal microscopy 48 h after *Agrobacterium* infiltration. From top to bottom: the GFP fluorescence spectrum of a representative nucleus from a transformed cell, an overlay of GFP and chlorophyll fluorescence and a bright field view of the scanned cell (Scale bars: 25 μm).

that the photosynthetic character is not a determining factor in this process. In agreement, an ORS278 mutant lacking the photosystem reaction center (Δ*pufLM*) nodulates *Aeschynomene sensitiva* roots as efficiently as the wild-type strain [49]. However, bacterial photosynthesis, gives these bacteria a functional advantage in stem nodulation, another rare property found in some *Aeschynomene* species, likely by contributing to the energetic requirement of nitrogenase and to the survival of the bacteria on the stem surface [49]. The molecular basis of this symbiotic process is still unknown but the recent identification of plant determinants involved in this symbiosis, with in particular the identification of a new putative membrane receptor [50], should allow to progress in the identification of the bacterial signal governing this symbiosis.

The NF-independent T3SS-dependent symbiosis, a more common process

Outside the photosynthetic supergroup of bradyrhizobia, all 59 strains inducing nodules or nodule-like structures on *A. indica* have in common that they harbor a *T3SS* gene cluster. Inactivating a key component (RhcN) in the T3SS machinery in representative strains resulted in all cases in a loss of nodulation confirming the key role of T3SS. More than 50% of the tested strains that have a T3SS are able to initiate a symbiosis. Nevertheless, none of them is capable of efficiently fixing nitrogen with *A. indica* (Supplementary Fig. S1). This does not seem to be correlated with the absence of the *nifHDK* genes, encoding the protein subunits of the nitrogenase enzyme complex, since these genes are present in all these strains, except for STM3557 (Fig. 2), but rather with an inability to differentiate properly into bacteroids, which corresponds to the active form of rhizobia capable of fixing nitrogen [39]. This raises the questions what is missing in these strains to be fully functional and if in nature a symbiosis solely dependent on a T3SS really exists? Indeed, it is intriguing that, except the Tv2a-2 strain, all the strains able to develop a T3SS-dependent symbiosis with *A. indica* also contain *nod* genes. These genes are dispensable in the context of *A. indica* but are necessary in almost all other *Bradyrhizobium*-legume interactions. So, what is the advantage for these bacteria to have both NFs and ET-Nods? Do these ET-Nods act synergistically with NFs to reinforce the nodulation efficiency of the strains on their natural NF-dependent host plants? Can the ET-Nods substitute for the NFs in case they are not properly recognized by the LysM-RLK plant receptors to enlarge the host spectrum of the bacteria?

Bradyrhizobium strains isolated from soybean nodules represent a high proportion of sequenced strains given the agronomic importance of this plant. Three-quarter of them (33/44) contain an ErnA homolog, while the other ET-Nods (Bel2-5, Sup3, Ubi1, Ub2 and Ubi3) are much less frequent (Fig. 2). Despite the high diversity of these soybean-nodulating strains, their ErnA proteins

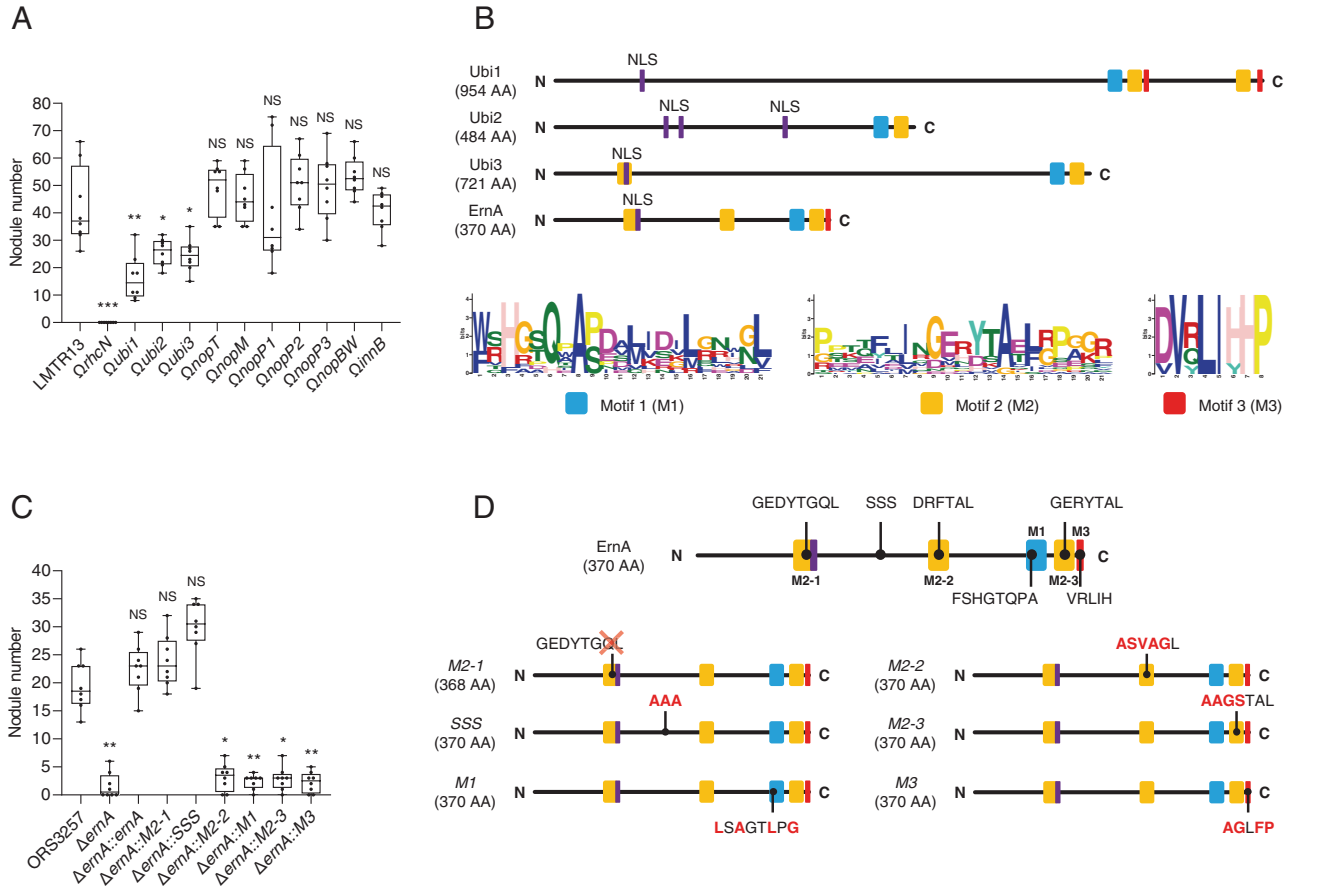


Fig. 8 The putative T3Es (Ubi1, Ubi2 and Ubi3) from *B. icense* LMTR13 share conserved motifs with ErnA that are important for nodulation in *A. indica*. **A** Nodule number on *A. indica* plants at 21 days after inoculation with strain LMTR13 and its mutant derivatives in T3SS (Ω*rhcN*) and various putative T3Es. Box plots show the results of one of the two experiments performed independently (8 plants each). *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$ significant differences between the wild-type strain and each mutant using a nonparametric Kruskal-Wallis test, NS: not significant. **B** Schematic representation of ErnA and the 3 new putative T3Es (Ubi1, Ubi2 and Ubi3). The conserved motifs (M1, M2 and M3) identified by MEME analysis (Fig. S4) are shown by colored squares and in web logos. The putative NLS (Nuclear localization signal) predicted by Localizer [43] are indicated by purple lines. **C** Nodule number on *A. indica* plants at 21 days after inoculation with ORS3257 and the ORS3257Δ*ernA* mutant expressing ErnA with mutations in the different identified motifs indicated in panel B. In the M2-1 motif, the mutagenesis was focused on the QL residues (deletion) considering that these residues are absent in the ErnA sequences clustering with $ErnA_{USDA110}$ shown to be less active in nodulation triggering on *A. indica* (Supplementary Fig. S2). Box plots show the results of one of the two experiments performed independently (8 plants each). *$p \leq 0.01$, **$p < 0.005$, significant differences between the wild-type strain and each mutant using a nonparametric Kruskal-Wallis test, NS: not significant. **D** Schematic representation of the different tested mutated forms of ErnA. The indicated SSS motif was identified as a putative DNA binding site using the software program DP Bind [45].

are strongly conserved (Supplementary Fig. S2). Although this ErnA variant seems less active in triggering nodulation in *A. indica* (Supplementary Fig. S2), we can assume that a strong selective pressure must be exerted on this ET-Nod during the *Bradyrhizobium-Glycine max* interaction to maintain such a high degree of conservation. ErnA and possible also other ET-Nods could therefore indeed play an important role in the fine tuning of the interaction between bradyrhizobia and their host plants, even if these strains dispose NFs to trigger nodulation.

Distinct ET-Nods induce nodule organogenesis

ErnA in strain ORS3257 is the first identified ET-Nod shown to activate the nodulation process in *A. indica* [12]. This T3E is conserved in more than two thirds of *Bradyrhizobium* strains with a T3SS (Fig. 2). The widespread ability among *Bradyrhizobium* strains to develop T3SS-dependent nodulation is most likely related to this high degree of conservation of this effector in strains having a T3SS. The loss of nodulation capacity observed with *ernA* mutants of various strains (ORS3257, USDA61) and the gain of nodulation observed after transferring ErnA in strain lacking it (DOA9) [12], support this hypothesis. However, this link between nodulation capacity and the presence of *ernA* is not strict as our study shows multiple examples of non-nodulating strains that harbor *ernA* and of strains that nodulate but do not contain *ernA*. Several non-exclusive reasons may explain the inability of some strains to nodulate *A. indica* while containing *ernA*: i) possessing a variant form of ErnA that would be less or not functional in the *A. indica* context, ii) the co-secretion of other T3Es triggering immune responses or iii) the presence or absence of other determinants involved in the recognition of the bacterium by the plant, such as EPS, KPS or LPS, known to modulate symbiotic and immune responses [51–53].

Among the strains able to nodulate in the absence of ErnA, we identified several new ET-Nod genes able to mimic the action of *ernA*. The first one, *sup3*, identified in several strains (NAS96.2, WSM1744 and ORS86) codes for a putative SUMO-protease. SUMO is a small protein related to the ubiquitin family that is used by eukaryotic cells to modify post-transcriptionally a large diversity of proteins [54]. The addition of the SUMO (SUMOylation process) or its cleavage by SUMO-proteases (deSUMOylation process) can

completely modify the biology of the targeted proteins, such as their localization, their interaction with other partner proteins, their enzymatic activity and their stability [55, 56]. SUMO-proteases are one of the most abundant classes of T3Es identified in bradyrhizobia [41]. Most of these SUMO-proteases, including Sup3 and Bel2-5, display the same architecture [56]. They correspond generally to very large proteins, often above one thousand AA, with a conserved C-terminal region corresponding to the SUMO-protease catalytic domain and a large N-terminal region containing no known functional domain. Phylogenetic analysis showed that the strains which have a SUMO-protease phylogenetically close to Sup3 in general nodulate *A. indica* and we confirmed in three of them (NAS96.2, WSM1744 and ORS86), the key role of Sup3 in controlling the nodulation process. In contrast, the SUMO-protease Bel2-5, which was found to be the key T3E regulating nodulation during the USDA61-*Glycine max* interaction, does not appear to play an important role in the context of *A. indica*. This indicates that there is a host plant-dependent specificity of nodulation activation of these SUMO proteases that is most likely related to their N-terminal region that is highly variable and supposed to be involved in substrate specificity [56]. Given the large diversity of SUMO-proteases present in bradyrhizobia, we expect that among them there are other ET-Nods able to activate nodulation on other host legumes.

How the different identified SUMO-proteases in *Bradyrhizobium* strains act to trigger the nodulation process remains unknown. A similar mechanism as used by the T3E XopD, present in some pathogenic bacteria of the genus *Xanthomonas*, which acts by deSUMOylation of a specific transcription factor (TF), can be assumed [57]. However, to our knowledge, none of the identified actors in legumes known to be recruited to the symbiotic signaling pathway [2, 50] have been reported to date to be modified by SUMOylation. To understand their mechanism in nodulation, future research should thus focus on the identification of the target(s) of these ET-Nod SUMO-proteases.

We observed that the addition of *sup3* to an *ernA* mutant can restore nodulation. This result is counterintuitive given that these two proteins are completely distinct. Two hypotheses can explain this result: i) these two ET-Nods trigger the same nodulation signaling pathway but acting on distinct actors at different levels of this pathway or, ii) these two ET-Nods could manipulate the same component(s) of this pathway but in different ways. Considering that the other putative ET-Nods (Ubi1, Ubi2 and Ubi3) identified in *B. jicamae* supergroup strains share certain motifs found in ErnA, this suggests a possible functional link between these different ET-Nods. In particular, a putative SIM domain was identified in Ubi1 and ErnA that appears to be essential for the nodulation induction activity of ErnA. This type of domain is found in many eukaryotic proteins that interact non-covalently with SUMO [47, 58]. All this would support the hypothesis that several identified ET-Nods have the same common SUMOylated target(s) but interact in a different way. For Sup3 and Bel2-5, this interaction would result in deSUMOylation of the target(s) and for ErnA and Ubi1 a modification of activity of the target(s) via a non-covalent protein-protein interaction thanks to the SUMO and the SIM domains. As the interaction of ErnA with SUMOs has to be validated and further functional studies are necessary to confirm that Ubi1, Ubi2 and Ubi3 are bona fide ET-Nods, this hypothesis remains speculative. However, this gives us some avenues to explore in future research to advance in the general understanding of the *modus operandi* of this newly discovered class of T3Es triggering nodulation.

In conclusion, this study provides much information on the factors predisposing *Bradyrhizobium* strains to develop a non-canonical symbiosis with legumes that is not based on NF perception. Many gaps remain, including the identification of the signals used by photosynthetic *Bradyrhizobium* strains to trigger nodulation and the understanding of the mode of action of the identified ET-Nods. It is important to continue this research on these alternative symbiotic systems to better estimate their occurrence and importance to trigger nodulation in nature, both in wild and crop legumes. Their understanding can shed new light on the evolution of symbiotic nitrogen-fixing mechanisms but also provide new levers to improve legume crop yields by optimizing their symbiotic capacity.

DATA AVAILABILITY

The genomic sequences of the 8 *Bradyrhizobium* strains sequenced in this study are available at NCBI BioProject (accession: PRJNA944954).

REFERENCES

1. Oldroyd GED. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol. 2013;11:252–63.
2. Roy S, Liu W, Nandety RS, Crook A, Mysore KS, Pislariu CI, et al. Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. Plant Cell. 2020;32:15–41.
3. Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E, Avarre JC, et al. Legumes symbioses: absence of *nod* genes in photosynthetic bradyrhizobia. Science. 2007;316:1307–12.
4. Okazaki S, Kaneko T, Sato S, Saeki K. Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. Proc Natl Acad Sci USA. 2013;110:17131–6.
5. Okazaki S, Tlttabutr P, Teulet A, Thouin J, Fardoux J, Chaintreuil C, et al. Rhizobium–legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. ISME J. 2016;10:64–74.
6. Mornico D, Miche L, Bena G, Nouwen N, Vermeglio A, Vallenet D. Comparative genomics of *Aeschynomene* symbionts: insights into the ecological lifestyle of *nod*-independent photosynthetic bradyrhizobia. Genes 2012;3:35–61.
7. Notti RQ, Stebbins CE. The structure and function of Type III secretion systems. Microbiol Spectr. 2016;4:VMBF-0004-2015. https://doi.org/10.1128/microbiolspec.
8. Hajra D, Nair AV, Chakravortty D. An elegant nano-injection machinery for sabotaging the host: Role of Type III secretion system in virulence of different human and animal pathogenic bacteria. Phys Life Rev. 2021;38:25–54.
9. Marie C, Broughton WJ, Deakin WJ. Rhizobium type III secretion systems: legume charmers or alarmers? Curr Opin Plant Biol. 2001;4:336–42.
10. Staehelin C, Krishnan HB. Nodulation outer proteins: double-edged swords of symbiotic rhizobia. Biochem J. 2015;470:263–74.
11. Teulet A, Camuel A, Perret X, Giraud E. The versatile roles of Type III secretion systems in rhizobium-legume symbioses. Annu Rev Microbiol. 2022;76:45–65.
12. Teulet A, Busset N, Fardoux J, Gully D, Chaintreuil C, CArtieaux F, et al. The rhizobial type III effector ErnA confers the ability to form nodules in legumes. Proc Natl Acad Sci USA. 2019;116:21758–68.
13. Ratu STN, Teulet A, Miwa H, Masuda S, Nguyen HP, Yasuda M, et al. Rhizobia use a pathogenic-like effector to hijack leguminous nodulation signaling. Sci Rep. 2021;11:2034. https://doi.org/10.1038/s41598-021-81598-6.
14. Ratu STN, Hirata A, Kalaw CO, Yasuda M, Tabuchi M, Okazaki S. Multiple domains in the rhizobial type III effector Bel2-5 determine symbiotic efficiency with soybean. Front Plant Sci. 2021;12:1–12.
15. Busset N, Gully D, Teulet A, Fardoux J, Camuel A, Cornu D, et al. The Type III effectome of the symbiotic *Bradyrhizobium vignae* strain ORS3257. Biomolecules. 2021;11:1592. https://doi.org/10.3390/biom11111592.
16. Vincent J. A manual for the practical study of root-nodule bacteria. Oxford, U.K: Blackwell Scientific Publications; 1970.
17. Sadowsky MJ, Tully RE, Cregan PB, Keyser HH. Genetic diversity in *Bradyrhizobium japonicum* serogroup 123 and its relation to genotype-specific nodulation of soybean. Appl Environ Microbiol. 1987;53:2624–30.
18. Gueye F, Moulin L, Sylla S, Ndoye I, Béna G. Genetic diversity and distribution of *Bradyrhizobium* and *Azorhizobium* strains associated with the herb legume *Zornia glochidiata* sampled from across Senegal. Syst Appl Microbiol. 2009;32:387–99.
19. Sambrook J, Fritsch EF, Maniatis TA. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory; 1989.
20. Plener L, Manfredi P, Valls M, Genin S. PrhG, a transcriptional regulator responding to growth conditions, is involved in the control of the type III secretion system regulon in *Ralstonia solanacearum*. J Bacteriol. 2010;192:1011–9.
21. Le Quéré A, Gully D, Teulet A, Navarro E, Gargani D, Fardoux J, et al. Complete genome sequence of *Bradyrhizobium* sp. strain ORS3257, an efficient nitrogen-fixing bacterium isolated from cowpea in Senegal. Microbiol Resour Announc. 2019;8:e01449-18. https://doi.org/10.1128/MRA.01449-18.

22. Orjuela J, Comte A, Ravel S, Charriat F, Vi T, Sabot F, et al. CulebrONT: a streamlined long reads multi-assembled pipeline for prokaryotic and eukaryotic genomes. Peer Comm J. 2022;2:e46. https://doi.org/10.24072/pcjournal.153.
23. Vallenet D, Belda E, Calteau A, Cruveiller S, Engelen S, Lajus A, et al. MicroScope-an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. Nucleic Acids Res. 2013;2013:636–47.
24. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.
25. Gautreau G, Bazin A, Gachet M, Planel R, Burlot L, Dubois M, et al. PPanGGOLiN: depicting microbial diversity via a partitioned pangenome graph. PLoS Comput Biol. 2020;16:e1007732. https://doi.org/10.1371/journal.pcbi.1007732.
26. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14:587–9.
27. Hoang D, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35:518–22.
28. Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49:W293–W296.
29. Sievers F, Barton GJ, Higgins DG. Multiple sequence alignment, vol 227. In: Baxevanis AD, Bader GD, Wishart DS, editors. Bioinformatics; Wiley; Hoboken, NJ, USA; 2020. p. 227–50.
30. Wongdee J, Songwattana P, Nouwen N, Noisangiam R, Fardoux J, Chaintreuil C, et al. *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. 2016;29:767–73.
31. Tsai JW, Alley MR. Proteolysis of the McpA chemoreceptor does not require the *Caulobacter* major chemotaxis operon. J Bacteriol. 2000;182:504–7.
32. Oke V, Long SR. Bacterial genes induced within the nodule during the Rhizobium-legume symbiosis. Mol Microbiol. 1999;32:837–49.
33. Bonaldi K, Gourion B, Fardoux J, Hannibal L, Cartieaux F, Boursot M, et al. 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. 2010;23:760–70.
34. Deslandes L, Olivier J, Peeters N, Feng DX, Khounlotham M, Boucher C, et al. Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. Proc Natl Acad Sci USA. 2003;100:8024–9.
35. Tasset C, Bernoux M, Jauneau A, Pouzet C, Brière C, Kieffer-Jacquinod S et al. Autoacetylation of the *Ralstonia solanacearum* effector PopP2 targets a lysine residue essential for RRS1-R-mediated immunity in *Arabidopsis*. PLoS Pathog. 2010. https://doi.org/10.1371/journal.ppat.1001202.
36. Lonjon F, Peeters N, Genin S, Vailleau F. In vitro and in vivo secretion/translocation assays to identify novel *Ralstonia solanacearum* Type 3 Effectors. Methods Mol Biol. 2018;1734:209–22.
37. Mattanovich D, Rüker F, da Camara Machado, Laimer M, Regner F, Steinkellner H, et al. Efficient transformation of *Agrobacterium tumefaciens* by electroporation. Nucleic Acids Res. 1989;17:6747.
38. Bonaldi K, Gherbi H, Franche C, Bastien G, Fardoux J, Barker D, et al. The Nod factor-independent symbiotic signaling pathway: development of *Agrobacterium rhizogenes*- mediated transformation for the legume *Aeschynomene indica*. Mol Plant Microbe Interact. 2010;23:1537–44.
39. 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.
40. Avontuur JR, Palmer M, Beukes CW, Chan WY, Coetzee MPA, Blom J, et al. Genome-informed *Bradyrhizobium* taxonomy: where to from here? Syst Appl Microbiol. 2019;42:427–39.
41. Teulet A, Gully D, Rouy Z, Camuel A, Koebnik R, Giraud E, et al. Phylogenetic distribution and evolutionary dynamics of *nod* and T3SS genes in the genus *Bradyrhizobium*. Micro Genom. 2020;6:mgen000407. https://doi.org/10.1099/mgen.0.000407.
42. Cunnac S, Occhialini A, Barberis P, Boucher C, Genin S. Inventory and functional analysis of the large Hrp regulon in *Ralstonia solanacearum:* identification of novel effector proteins translocated to plant host cells through the type III secretion system. Mol Microbiol. 2004;53:115–28.
43. Sperschneider J, Catanzariti AM, DeBoer K, Petre B, Gardiner DM, Singh KB, et al. LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. Sci Rep. 2017;7:44598. https://doi.org/10.1038/srep44598.
44. Bailey TL, Johnson J, Grant CE, Noble WS. "The MEME Suite". Nucleic Acids Res. 2015;43:W39–W49.
45. Hwang S, Gou G, Kuznetsov IB. DP-Bind: a web server for sequence-based prediction of DNA-binding residues in DNA-binding proteins. Bioinformatics 2007;23:634–6.
46. Zhao Q, Xie Y, Zheng Y, Jiang S, Liu W, Mu W, et al. GPS-SUMO: A tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Res. 2014;42:W325–330. https://doi.org/10.1093/nar/gku383.
47. Kerscher O. SUMO junction-what's your function? New insights through SUMO-interacting motifs. EMBO Rep. 2007;8:550–5. https://doi.org/10.1038/sj.embor.7400980.
48. Miché L, Moulin L, Chaintreuil C, Contreras-Jimenez JL, Munive-Hernández JA, Del Carmen Villegas-Hernandez M, et al. 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. 2010;12:2152–64.
49. 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. 2000;26:14795–800.
50. Quilbé J, Lamy L, Brottier L, Leleux P, Fardoux J, Rivallan R, et al. Genetics of nodulation in *Aeschynomene evenia* uncovers mechanisms of the rhizobium-legume symbiosis. Nat Commun. 2021;121:829. https://doi.org/10.1038/s41467-021-21094-7.
51. Gibson KE, Kobayashi H, Walker GC. Molecular determinants of a symbiotic chronic infection. Annu Rev Genet. 2008;42:413–41.
52. Janczarek M, Rachwał K, Marzec A, Grządziel J, Palusińska-Szysz M. Signal molecules and cell-surface components involved in early stages of the legume-rhizobium interactions. Appl Soil Ecol. 2015;85:94–113.
53. Poole P, Ramachandran V, Terpolilli J. Rhizobia: from saprophytes to endosymbionts. Nat Rev Microbiol. 2018;16:291–303.
54. Johnson ES. Protein modification by SUMO. Ann Rev Biochem. 2004;73:355–82.
55. Augustine RC, Vierstra RD. SUMOylation: re-wiring the plant nucleus during stress development. Curr Opin Plant Biol. 2018;45:143–54.
56. Morrell R, Sadanandom A. Dealing with stress: a review of plant SUMO proteases. Front Plant Sci. 2019;10:1122.
57. Kim JG, Stork W, Mudgett MB. *Xanthomonas* type III effector XopD desumoylates tomato transcription factor SlERF4 to suppress ethylene responses and promote pathogen growth. Cell Host Microbe. 2013;13:143–54.
58. Yau TY, Sander W, Eidson C, Courey AJ. SUMO interacting motifs: structure and function. cells. 2021;10:2825. https://doi.org/10.3390/cells10112825.
59. Tria FDK, Landan G, Dagan T. Phylogenetic rooting using minimal ancestor deviation. Nat Ecol Evol. 2017;1:193. https://doi.org/10.1038/s41559-017-0193.
60. Sanchez C, Mercante V, Babuin MF, Lepek VC. Dual effect of *Mesorhizobium loti* T3SS functionality on the symbiotic process. FEMS Microbiol Lett. 2012;330:148–56.
61. Rodrigues JA, Lopez-Baena FJ, Ollero FJ, Vinardell JM, Espuny MR, Bellogin RA, et al. NopM and NopD are rhizobial nodulation outer proteins: identification using LC-MALDI and LC-ESI with a monolithic capillary column. J Proteome Res. 2007;6:1029–37.
62. Okazaki S, Zehner S, Hempel J, Lang K, Göttfert M. Genetic organization and functional analysis of the type III secretion system of *Bradyrhizobium elkanii*. FEMS Microbiol Lett. 2009;295:88–95.
63. Faruque OM, Miwa H, Yasuda M, Fujii Y, Kaneko T, Sato S, et al. Identification of *Bradyrhizobium elkanii* genes involved in incompatibility with soybean plants carrying the Rj4 Allele. Appl Environ Microbiol. 2015;81:6710–7.

ACKNOWLEDGEMENTS

This work was supported by a grant from the French National Research Agency ("ET-Nod"; ANR-20-CE20-0012), the French Laboratory of Excellence project "TULIP" (ANR-10-LABX-41; ANR-11-IDEX-0002-02) and the "Ecole Universitaire de Recherche (EUR)" TULIP-GS (ANR-18-EURE-0019). AC was supported by a PhD fellowship from the French Ministry of National Education, Higher Education and Research. We thank all the researchers who sent us their *Bradyrhizobium* strains: Neung Teaumroong, Panlada Tittabutr, Kiwamu Minamisawa, Xinhua Sui, Christian Staehelin, Fabricio Cassan, Pekka Oivanen, Penny Hirsch, Kristen DeAngelis, Eric Roden, Els Vercoutere, Nadine Tollens, Hirohito Tsurumaru, Julia Vorholt, William Mohn, Luciane Passaglia, Dale Pelletier, Tania Wade, Jinjin Tao, Stephanus Venter, Ederson da Conceição Jesus, Eliana Gertrudes de Macedo Lemos, Matthew Parker, François Lefort, Djamel Chekireb, Eoin Brodie, Hans Martin Fischer, Mike Sadowsky, John Ferguson.

AUTHOR CONTRIBUTIONS

AC, AT, PM, LD, and EG conceived and designed the study. AC, AT, MC, FH, VP, DG, MP, CC, JF, NHA, SO, STNR, FG, JZ, NN, JFA, HL, PM, LD, and EG performed the experiments and analyzed the results. AC, LD, PM and EG wrote the manuscript. All authors commented on the results and the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41396-023-01458-1.

Correspondence and requests for materials should be addressed to Eric Giraud.

Reprints and permission information is available at http://www.nature.com/reprints

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023

Supplementary information for

Identification in the *Bradyrhizobium* genus of diverse Type 3 effectors triggering legume nodulation in the absence of Nod factors

Alicia Camuel[1,2], Albin Teulet[1,3], Mélanie Carcagno[4], Fazal Haq[5], Valérie Pacquit[4], Djamel Gully[1,2], Marjorie Pervent[1,2], Clémence Chaintreuil[1,2], Joël Fardoux[1], Natasha Horta-Araujo[1,2], Shin Okazaki[6], Safirah Tasa Nerves Ratu[6], Fatou Gueye[7], Jerri Zilli[8], Nico Nouwen[1,2], Jean-François Arrighi[1,2], Haiwei Luo[9], Peter Mergaert[5], Laurent Deslandes[4], Eric Giraud[1,2,*]

* Corresponding author: eric.giraud@ird.fr

This document includes only Figs. S1 to S4.

Datasets S1 to S5 are available on: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10432411/

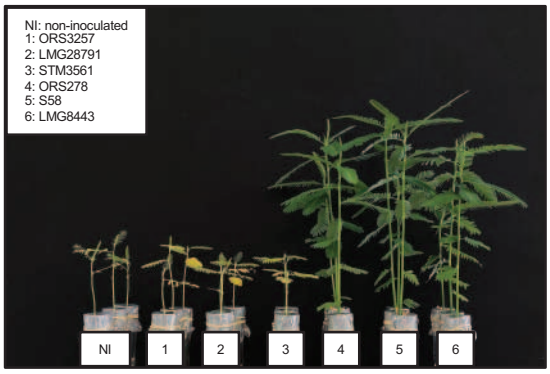


Figure S1. The NF-independent T3SS-dependent symbiosis is less efficient for fixing nitrogen than the NF-independent T3SS-independent symbiosis in *A. indica*.
Comparison of the growth of the plant (leaf phenotype) non-inoculated (NI), inoculated with strains able to fix nitrogen with a low efficiency in a T3SS-dependent manner (Type E): ORS3257 (1), LMG28791 (2) and STM3561 (3) ; or with strains belonging to photosynthetic supergroup able to fix efficiently nitrogen in a T3SS-independent manner (Type F): ORS278 (4), S58 (5) and LMG8443 (6). Pictures were taken 21 days after inoculation.

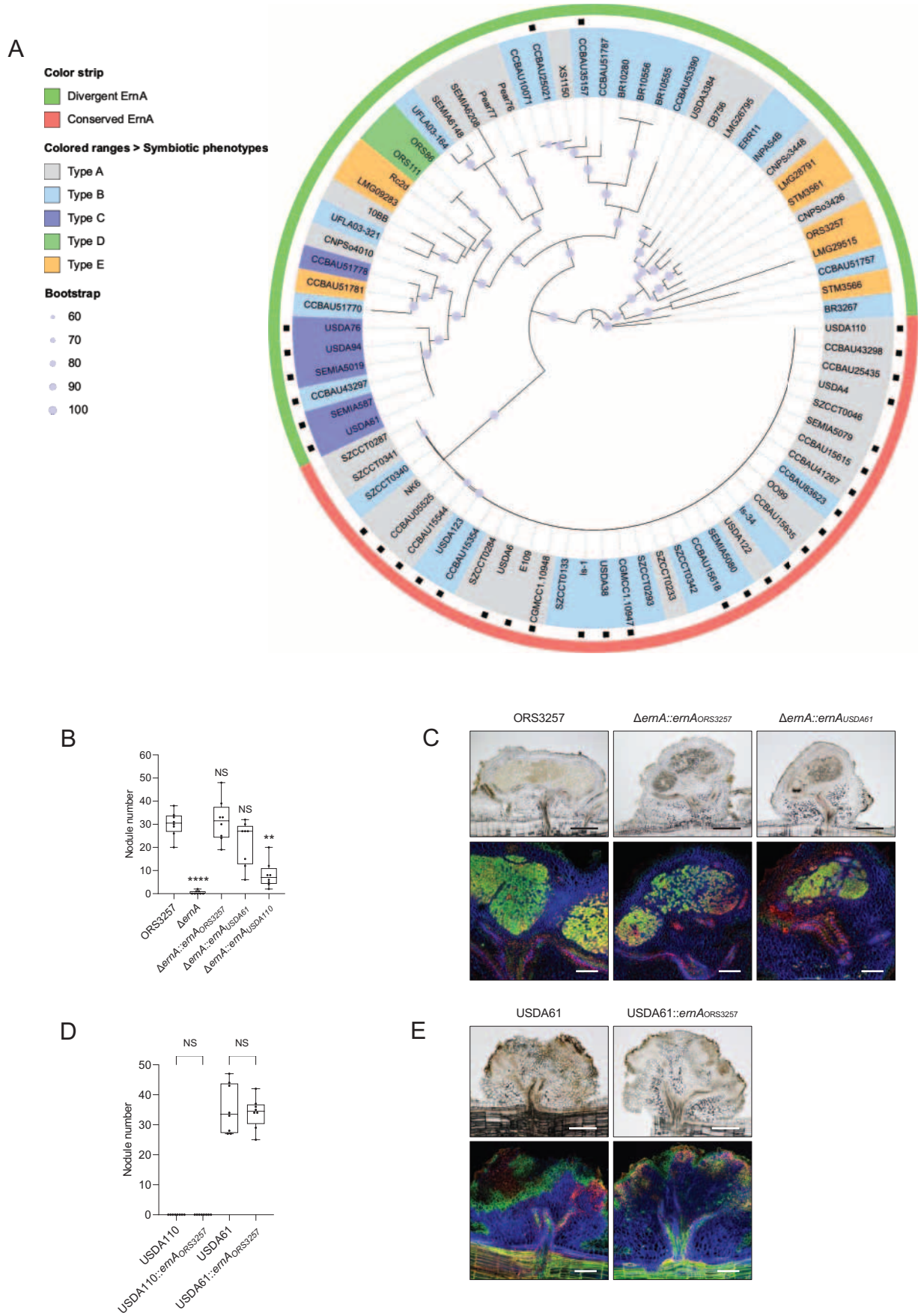


Figure S2. Exploring the natural variability of ErnA on nodulation capacity and infection.
(A). Phylogenetic tree of full-length ErnA homologs identified in the collection of bradyrhizobia used in this study. Two main clusters are identified and indicated by colour strip according to their level of conservation. The symbiotic phenotype of each strain having an ErnA homolog is indicated by a coloured range referenced in the key. The black squares indicate strains isolated from soybean nodule. (B and D). Nodule number on *A. indica* plants at 21 days after inoculation with ORS3257 (B) or USDA61 and USDA110 (D) and their respective mutants containing various *ernA* genes. Box plots show the results of one of the two experiments performed independently (8 plants each). ** $p \leq 0.01$, **** $p < 0.0001$ significant differences between the wild-type strain ORS3257 and each mutant using a parametric Welch test, NS: not significant. No significant differences were calculated between the wild-type strain USDA61 and USDA110 with their respective mutant (D) using a nonparametric Kruskal-Wallis test, NS: not significant. (C and E) Cross-sections of nodules were observed by light or confocal microscopy after staining with SYTO 9, propidium iodide and calcofluor; scale bars C and E, 200 µm. Intracellular infection is visible in C while only intercellular infection can sometimes be observed in E.

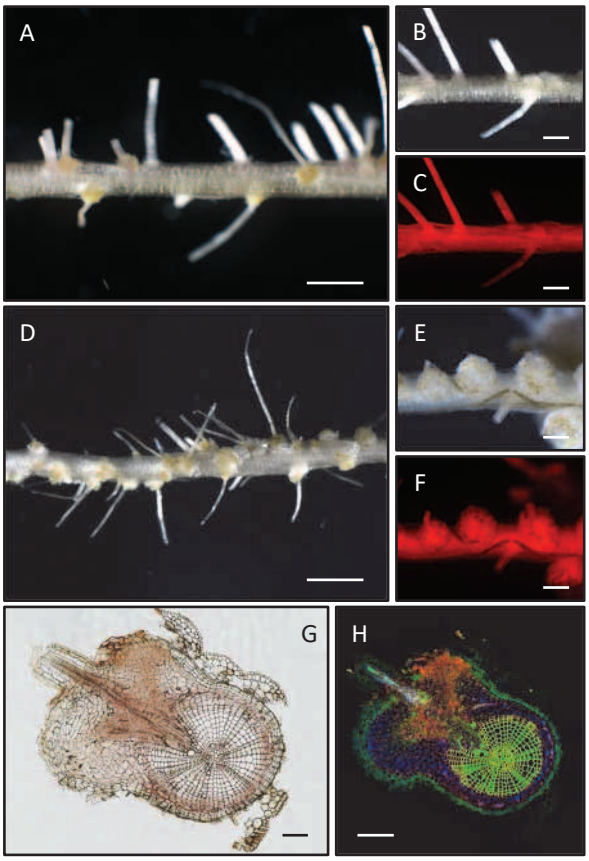
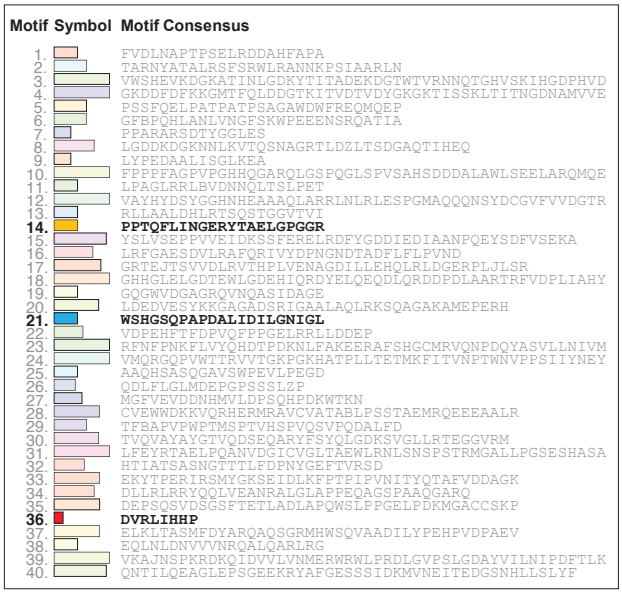


Figure S3. Ectopic expression of *sup3* in *A. indica* roots induces spontaneous nodules.
(A-F) *A. indica* roots transformed either with the empty vector containing the DsRed marker (A-C) or with *p35S-sup3* (D-F) at 7 weeks after transformation and in the absence of bradyrhizobia. Roots were observed by a light and fluorescence stereomicroscope equipped with a DsRed filter. (Scale bars: A and D, 2 mm; B, C, E and F, 500 µm). Pseudo-nodules have peripheral vascularisation (G and H). Cross-sections of pseudo-nodules were observed by light or confocal microscopy after staining with SYTO 9, propidium iodide and calcofluor; scale bars G and H, 200 µm. No bacteria were observed in the root tissues.

A

B



C

Sup3_WSM1744_JAAVLW01_10	Sup3_NAS96.2_LGHK01_v1_3
Bel2-5_USDA61_BE61_51970	ErnA_ORS3257_BRAD3257_v2
NopE_LMTR13_24220	NopH_LMTR13_24225
NopC_LMTR13_24215	NopL_LMTR13_24195
NopM_LMTR13_24790	NopP1_LMTR13_24310
NopP2_LMTR13_24260	NopP3_LMTR13_24320
NopT_LMTR13_23985	NopAF_LMTR13_04200
NopAJ_LMTR13AM_6340	NopBW_LMTR13_24270
InnB_LMTR13_24100	Ubi1_LMTR13_24185
Ubi2_LMTR13AM_5744	Ubi3_LMTR13_27190
NopH_USDA110_bll1804	NopE1_USDA110_blr1806
NopE2_USDA110_blr1649	NopM_USDA110_blr1676
NopP_USDA110_blr1752	NopT1_USDA110_blr2058
NopT2_USD110_blr2140	NopAA_USDA110_blr1656
NopAB_USDA110_blr1869	NopAC_USDA110_blr1993
NopAF_USDA110_blr7131	NopAG_USDA110_bll1862
NopAK_USDA110_blr1789	NopAJ_USDA110_bll1846
NopAH_USDA110_bll1798	NopAI_USDA110_bll1796
NopAN_USDA110_blr1912	NopAR_USDA110_bll1840
NopAS_USDA110_blr1626	NopAZ_USDA110_blr0077
NopBA_USDA110_blr1787	NopBK_USDA110_blr1975
NopX_S.fredii_NGR_a00700	NopBW_M.loti_mlr6358
NopL_S.fredii_NGR_a00770	NopJ_S.fredii_NGR_a02610
NopC_S.fredii_AAY33494	InnB-USDA61_BE61_78180
NopD_S.fredii_CEO91485	NopBX_M.loti_mlr6361

Figure S4. MEME analysis on putative ET-Nods and functionally characterized T3Es.
(A). Schematic representation of the conserved motifs (40) identified in an input of 50 rhizobial T3Es using the MEME suite tools [44]. This input includes the 13 putative T3Es identified in LMTR13 and 37 rhizobial T3Es previously characterized [10,12]. The parameters used for the MEME analysis were the following : i) The maximum number of motif identification was restricted to 40, ii) The occurrence of motif was set to any number of repetitions and iii) the minimum motif width was set to 6 and a maximum of 50. The panel A displays only the 35 T3Es in which conserved motives were identified by MEME. The ID of the protein, the p-value and the motif locations identified into the protein appears in column 1, 2 and 3 respectively. ErnA, Ubi1, Ubi2 and Ubi3 are shown in bold. (B). Corresponding colour and consensus of each motif. M1, M2 and M3 conserved motifs appear in bold, respectively in the 14th, 21st and 36th position. (C). Complete list of the T3Es used as input.

Chapitre 2

1. Contexte de l'étude

Un des objectifs sous-jacents à la caractérisation de la capacité des *Bradyrhizobium* à utiliser un processus alternatif indépendant des FNs pour déclencher la nodulation a été d'initier le décryptage de la voie de signalisation symbiotique activée par les ET-Nods.

Il a été précédemment montré qu'*A. evenia* est une espèce pertinente pour la caractérisation fonctionnelle du processus Nod-indépendant, T3SS-indépendant médié par les *Bradyrhizobium* photosynthétiques (Arrighi & Cartieaux, 2015), et son génome est séquencé (Quilbé, Nouwen, et al., 2022). Cependant, la lignée utilisée (N76) est très mal nodulée par les souches de *Bradyrhizobium* non-photosynthétiques utilisant leur T3SS pour induire la nodulation. Aussi, nous avons recherché parmi les multiples lignées disponibles d'*A. evenia* si certaines étaient mieux nodulées par ce type de processus symbiotique dépendant du T3SS. Parmi la vingtaine de lignées d'*A. evenia* testées avec plusieurs souches de *Bradyrhizobium*, la lignée 21 (N21) est apparue comme celle répondant le mieux au processus T3SS-dépendant. Il a ainsi été montré que les réponses symbiotiques initialement observées sur *A. indica* sont conservées chez *A. evenia* N21 avec les principales souches qui ont précédemment émergées (Camuel et al., 2023). L'ensemble de ces observations indiquent qu'*A. evenia* N21 est un bon modèle pour l'étude fonctionnelle de la symbiose Nod-indépendante, T3SS-dépendante.

Dans le but de mieux comprendre le mode d'action des ET-Nods et d'identifier des cibles potentielles d'ErnA, il a été mené une analyse RNA-seq de plantes d'*A. evenia* N21 inoculées par la souche sauvage *B. vignae* ORS3257 ou par le mutant de délétion ORS3257Δ*ernA*. Les données transcriptomiques obtenues montrent en particulier une augmentation de l'expression de certains gènes de défense de la plante ou impliqués dans la signalisation de l'auxine. Cette analyse nous donne ainsi des premières pistes sur les potentielles cibles d'ErnA et de son mode d'action.

Par ailleurs, bien que la lignée N76 ne soit pas idéale pour l'étude de la symbiose T3SS-dépendante, nous avons voulu mettre à profit les ressources génétiques disponibles au sein de l'équipe MSLT sur celle-ci. En effet, plusieurs mutants EMS dans des gènes symbiotiques connus (*AePOLLUX*, *AeCCaMK*, *AeCYCLOPS*, *AeNSP2* et *AeNIN*) ou nouveau (*AeCRK*), ont pu être sélectionnés pour leur incapacité à être nodulés par la souche photosynthétique *Bradyrhizobium* sp. ORS278 (T3SS-indépendante). Ces mutants apparaissent ainsi propices à déterminer si ces gènes sont requis également dans la voie de signalisation T3SS-dépendante et si des différences peuvent être observées selon l'ET-Nod injecté. Cinq souches demeurant capables de noduler la lignée sauvage d'*A. evenia* N76 et représentant le mieux possible la diversité du genre *Bradyrhizobium* et des ET-Nods jusqu'alors identifiés ont ainsi

été sélectionnées. Des différences et des points communs dans la capacité à induire des nodules ont été observés selon les souches et les mutants de plantes testés. De ces analyses, il ressort que les divers ET-Nods identifiés pourraient intervenir à différents niveaux d'une voie de signalisation déclenchant la nodulation et qu'ainsi certains déterminants symbiotiques clés (AePOLLUX, AeCCaMK, AeCYCLOPS) qui ont été montrés jusqu'alors requis pour plusieurs symbioses (nodulation Nod-dépendante, Nod-indépendante, T3SS-indépendante et mycorhization) pourraient être dispensables dans le processus T3SS-dépendant suivant le ou les effecteur(s) injecté(s). Par opposition, la mise en évidence que NIN et NSP2 demeurent des acteurs requis quelle que soit la souche testée vient appuyer leur rôle de master-régulateur dans les diverses symbioses nodulaires existantes.

Enfin, deux autres points notables émergent de cette étude : i) il existe une coopérativité de certains ET-Nod (ErnA et Sup3) qui permet de court-circuiter certains déterminants symbiotiques majeurs tels que POLLUX, CCaMK et CYCLOPS, et ii) CRK, précédemment identifié comme un acteur clé de la symbiose Nod-indépendante, T3SS-indépendante, apparaît dispensable chez l'ensemble des souches testées utilisant leur T3SS pour déclencher la nodulation, ce qui renforce l'idée que ce déterminant serait spécifique à la symbiose T3SS-indépendante chez *Aeschynomene* et reconnaîtrait le ou les signaux "non Nod".

L'ensemble de cette étude a permis de faire émerger un nouveau modèle de plante, *A. evenia* N21, pour l'étude du processus Nod-indépendant, T3SS-dépendant. Ces travaux ouvrent ainsi la voie à de nombreuses approches fonctionnelles pour mieux déchiffrer la voie de signalisation requise pour cette symbiose et comprendre le mode d'action des ET-Nods.

2. Article 2 : "New insights into the T3SS-dependent symbiotic pathway between *Aeschynomene evenia* and *Bradyrhizobium* strains"

New insights into the T3SS-dependent symbiotic pathway between *Aeschynomene evenia* and *Bradyrhizobium* strains

Alicia Camuel[1,2], Djamel Gully[1,2], Marjorie Pervent[1,2], Albin Teulet[1,3], Nico Nouwen[1,2], Jean-François Arrighi[1,2*§], Eric Giraud[1,2,*§]

1- IRD, Laboratoire des Symbioses Tropicales et Méditerranéennes (LSTM), UMR IRD/ Institut Agro /INRAE/ Université de Montpellier /CIRAD, TA-A82/J– Campus de Baillarguet 34398 Montpellier cedex 5, France.

2-PHIM Plant Health Institute, Université de Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France.

3-University of Cambridge, Sainsbury Laboratory (SLCU), Cambridge, CB2 1LR, UK

§ : these authors have contributed equally to this work

* Corresponding authors: eric.giraud@ird.fr, jean-francois.arrighi@ird.fr

Competing of interest statement: The authors declare that there are any competing financial interests in relation to the work describe

Abstract

Some *Aeschynomene* species display the original property of being nodulated independently of Nod factors (NFs) by two distinct processes that differ in the requirement for a type III secretion system (T3SS). While different T3 effectors (T3Es) triggering nodulation (ErnA and Sup3) were identified, nothing is known about the signaling pathways they activate. In this study, by exploring intraspecific variation in nodulation capacity of the model species *Aeschynomene evenia*, we identified the accession N21/PI 225551 as appropriate to decipher the NF/T3SS-dependent process. Using this accession, comparative transcriptomic analysis of roots inoculated with the ORS3257 strain and its Δ*ernA* mutant was performed and revealed a panel of genes differentially expressed including genes involved in plant defenses and Auxin signaling. Furthermore, we analyzed the symbiotic phenotype of 5 *Bradyrhizobium* strains having different sets of T3Es on *A. evenia* N76/CIAT22838 plants with mutations in the known symbiotic genes *AeCRK*, *AePOLLUX*, *AeCCaMK*, *AeCYCLOPS*, *AeNSP2* and *AeNIN*. While mutation in *AeCRK* had no effect on nodulation, all the strains were unable to nodulate *AeNIN* and Ae*NSP2* mutant lines, indicating that these genes are required for T3SS-dependent nodulation. The nodulation phenotype of *AePOLLUX*, *AeCCaMK* and AeCYCLOPS mutant lines greatly depended on the strains tested. Notably, the ORS86 strain can nodulate these 3 mutant lines and requires for this the cooperation of both ErnA and Sup3. Overall, this study improves our knowledge on how the T3SS-dependent nodulation process is achieved in legumes.

Introduction

In rhizobium/legume symbiosis, the interaction leads to the formation of a symbiotic organ, the nodule, in which the bacteria fix dinitrogen for the plant benefit. This interaction generally involves an exchange of diffusible signals between the two partners and in which the rhizobial Nod factors (NFs) are recognized as the key signal that governs the nodule formation and its concomitant infection by the bacteria (Perret et al. 2000; Oldroyd 2013). However, alternative symbiotic processes independent of the NF signal have also been described (Giraud et al. 2007; Okazaki et al. 2013 and 2016). For instance, some tropical legumes of the *Aeschynomene* genus, such as *Aeschynomene indica* and *Aeschynomene evenia,* are nodulated by photosynthetic *Bradyrhizobium* strains (ORS278 and BTAi1) lacking the canonical *nodABC* genes necessary for NF synthesis (Giraud et al. 2007). The bacterial determinant(s) that trigger(s) this NF-independent nodulation process still remain unknown. Interestingly, in *in vitro* conditions, *A. indica* can also be nodulated by a large diversity of non-photosynthetic bradyrhizobia (Okazaki et al. 2016; Camuel et al. 2023). Contrary to the photosynthetic strains, this interaction is possible only if the bacteria have a functional type III secretion system (T3SS). This secretory machinery is capable to inject bacterial proteins, named Type III effectors (T3Es), directly into the host cell, thereby modulating the host immunity and inducing nodule organogenesis (Teulet et al. 2022). In the case of the T3SS-dependent symbiosis between *B. vignae* ORS3257 strain and *A. indica*, it has been shown that at least 5 T3Es are required: NopM and NopP repress the plant defense immunity, NopT and NopAB are required for bacterial infection, and finally ErnA activates nodule organogenesis (Teulet et al. 2019). Interestingly, with other *Bradyrhizobium* strains, *A. indica* nodule organogenesis is activated by other T3ES, such as Sup3 encoding a putative SUMO protease in ORS86 and WSM1744 strains, and 3 unknow putative T3Es (named Ubi1, Ubi2 and Ubi3) in the LMTR13 strain (Camuel et al. 2023). These various T3Es, proposed to be called ET-Nods for effectors triggering nodulation, are very different from each

other suggesting variations in the *modus operandi* to activate the T3SS-dependent nodulation program in *Aeschynomene*.

During the last two decades, very important progress has been made to identify the plant genetic determinants that are involved in the Nod factor signaling pathway. Using *Medicago truncatula* and *Lotus japonicus* as model plants, the main actors governing symbiosis have been identified (Roy et al. 2020). In brief, upon NFs perception, a pair of LysM-RLK receptors (MtNFP/LjNFR5 – MtLYK3/LjNFR1) and the co-receptor SYMRK (MtDMI2) activate nuclear-associated calcium oscillations via several cation channels (CASTOR/MtDMI2 and POLLUX/MtCNGC15) and nucleoprorins (LjNUP85, LjNUP133 and NENA). A Ca^{2+}/calmodulin-dependent kinase (CCaMK/MtDMI3) decodes these calcium oscillations and phosphorylates the transcription factor CYCLOPS/IPD3 which, together with other transcription factors (NSP1 and NSP2), activates the expression of NIN, the master regulator of nodule initiation and infection thread development (Feng et al. 2021).

To advance in the understanding of the NF/T3SS-independent symbiosis, *A. evenia* was chosen to develop functional studies (Arrighi et al. 2012). Its small diploid genome size (2n=20, 415 Mb/1C), its selfing nature and its prolific seeds production were strong criteria to select this species as model. The *A. evenia* inbred CIAT22838 line was used to establish a reference genome sequence and to develop a forward genetic approach. By screening 70,000 ethyl methanesulfonate (EMS)-mutagenized seedlings for defects in nodulation with the photosynthetic *Bradyrhizobium* model strain ORS278, several mutant lines with a complete absence of nodulation were identified and shown to have mutations in key determinants (*AePOLLUX*, *AeCCaMK*, *AeCYCLOPS*, *AeNSP2* and *AeNIN*) of the NF-signaling pathway (Quilbé et al. 2021 and 2022). This indicates that the NF/T3SS-independent signaling process recruits several conserved symbiotic determinants acting downstream of the NF-receptors. Interestingly, this mutant-based approach also identified a new symbiotic gene *AeCRK*

encoding a cysteine-rich receptor kinase (CRK) that may play a key role in sensing the non-Nod signal produced by photosynthetic *Bradyrhzobium* strains to activate NF-independent symbiosis (Quilbé et al. 2021 and 2022).

In contrast to the two above described symbiotic processes, our knowledge of the NF/T3SS-dependent process remains extremely limited. It is hypothesized that ErnA, which has been shown to target the host cell nucleus and interact with nucleic acids, may act by modulating plant genes expression, while Sup3, also targeted to the host nucleus but containing a C-terminal small ubiquitin-like modifier (SUMO)-Protease domain, may act by deSUMOylation of specific nuclear plant protein targets (Teulet et al. 2019; Camuel et al. 2023). Thus, these 2 distinct ET-Nods might target different actors to trigger nodulation, which might furthermore not correspond to the conserved symbiotic determinants shown to be required for the NF-dependent and NF/T3SS-independent signaling pathways. Furthermore, within *A. evenia,* a high genetic diversity was observed, which in contrast to nodulation with photosynthetic *Bradyrhizobium* strains, has a strong impact on nodulation by *Bradyrhizobium* strains using a NF/T3SS-dependent process (Chaintreuil et al. 2018; Okazaki et al. 2016). This indicates that within *A. evenia,* intraspecific variations of some specific determinant(s) govern the outcome of the symbiotic interaction when the T3SS is involved. Notably, the *A. evenia* CIAT22838 accession that is used as model plant to decipher the molecular mechanism of the NF/ T3SS-independent process is not or very badly nodulated by strains using a NF/T3SS-dependent process (Okazaki et al. 2016).

In this study, to further our understanding of the NF/T3SS-dependent symbiosis, we have explored the intraspecies variability in nodulation within *A. evenia* and identified an appropriate line (PI 225551, named here N21) that is well nodulated by non-photosynthetic *Bradyrhizobium* strains using their T3SS to trigger nodulation. Using *A. evenia* N21, we performed a comparative RNASeq analysis of roots inoculated with ORS3257 and its *ernA*

mutant to identify the host-plant genes whose expression is impacted by ErnA. Finally, we exploited the *A. evenia* CIAT22838 (referred here as N76) mutant lines impaired in nodulation with the ORS278 strain, to determine whether some non-photosynthetic *Bradyrhizobium* strains can bypass, thanks to their T3SS, the requirement of certain canonical symbiotic genes for nodulation.

Materials and Methods

Plant material, growth conditions

All accessions of *Aeschynomene* used in this study are detailed in Supplementary Table S1. The *A. evenia* mutant lines affected in the common symbiotic pathway genes were obtained from the phenotypic screen of an EMS-mutagenized population derived from the reference CIAT22838 line (Quilbé et al. 2021). Seeds were scarified for 40 min with sulfuric acid (96% v/v) and rinsed with distilled water. Germination was induced overnight at room temperature in distilled water for wild-type seeds, and with 0.01% (v/v) of ethrel (BAYER) for mutant's seeds. Germinated seeds were transferred onto 0.8% water agar plates and incubated overnight at 34 °C. *A. evenia* and *A. indica* seedlings were cultured in cultivation chamber as previously described (Arrighi et al, 2012).

Bacterial strains and growth conditions

All bacterial strains and derivative mutants used are listed in Supplementary Table S2. The *Bradyrhizobium* strains were grown in YM or AG medium at 28 °C (Vincent, 1970; Sadowsky et al. 1987). *Agrobacterium rhizogenes* Arqua1 strain was grown at 28 °C on AG medium. *Escherichia coli* strains were grown at 37 °C in modified Lysogeny Broth (LB) medium (Sambrook et al. 1989). When required, the media were supplemented with the appropriate antibiotics at the following concentrations: 20 µg/mL nalidixic acid, 20 µg/mL cefotaxime, 50-100 µg/mL kanamycin and 100-200 µg/mL spectinomycin.

***Sup3* transfer into ORS3257 strain.**

The complete *sup3* (JAAVLW01_10281) gene of WSM1744 was amplified by PCR using the Phusion™ High-Fidelity DNA Polymerase (Thermo Scientific) and the primers (GGCCTAT**ACTAGT**ACCGAGGGGGACATGAAGTTCCAATCCACCAACTG / CAGCCAA**ACTAGT**TTAGCCCAAGCCGGGATGAGCCAGTCGACTC). The PCR product was cloned into the *Spe*I-site of plasmid pVO155-*pm-ernA* (Camuel et al. 2023) which places the *sup3* gene under the control of the *ernA* ORS3257 promoter. The construct was transferred into the ORS3257 strain by single crossing over as previously described (Teulet et al. 2019).

Symbiotic analysis

For each condition, eight to ten plants were inoculated with a bacterial strain and symbiotic properties (number of nodules per plant / nitrogenase enzyme activity) were analysed 21 days after inoculation (Bonaldi et al. 2010a). Cytological analysis of the nodules was performed as described (Songwattana et al. 2021).

RNA-seq analysis

A. evenia N21 plants were inoculated with the ORS3257 WT strain and its Δ*ernA* mutant. At 2, 3, 4 and 9 days after inoculation, the root system of 12 plants per conditions were pooled, immediately frozen in liquid nitrogen and subsequently grounded into a fine powder. For each time-point, three replicates were done. For total RNA extraction, 100 mg of the powder was taken and used in the protocol as described by Chomczynski *et al*., 1987. RNA quality and concentration were evaluated using a NanoDrop 2000/200c (Thermo Fisher) and an Agilent 2100 Bioanalyzer according to the manufacturer's instructions. The cDNA library was constructed as previously described (Gully et al., 2018). cDNA libraries were sequenced on a full S1 flow-cell (NovaSeq) in single read 100 nt mode. Image analysis, base calling and quality filtering were performed by MGX platform (Montpellier GenomiX). The 'raw reads' were filtered to remove low quality and contaminant reads, including adaptor sequences, DNA, or PCR duplicates, which generated trimmed reads. The trimmed reads were mapped to the

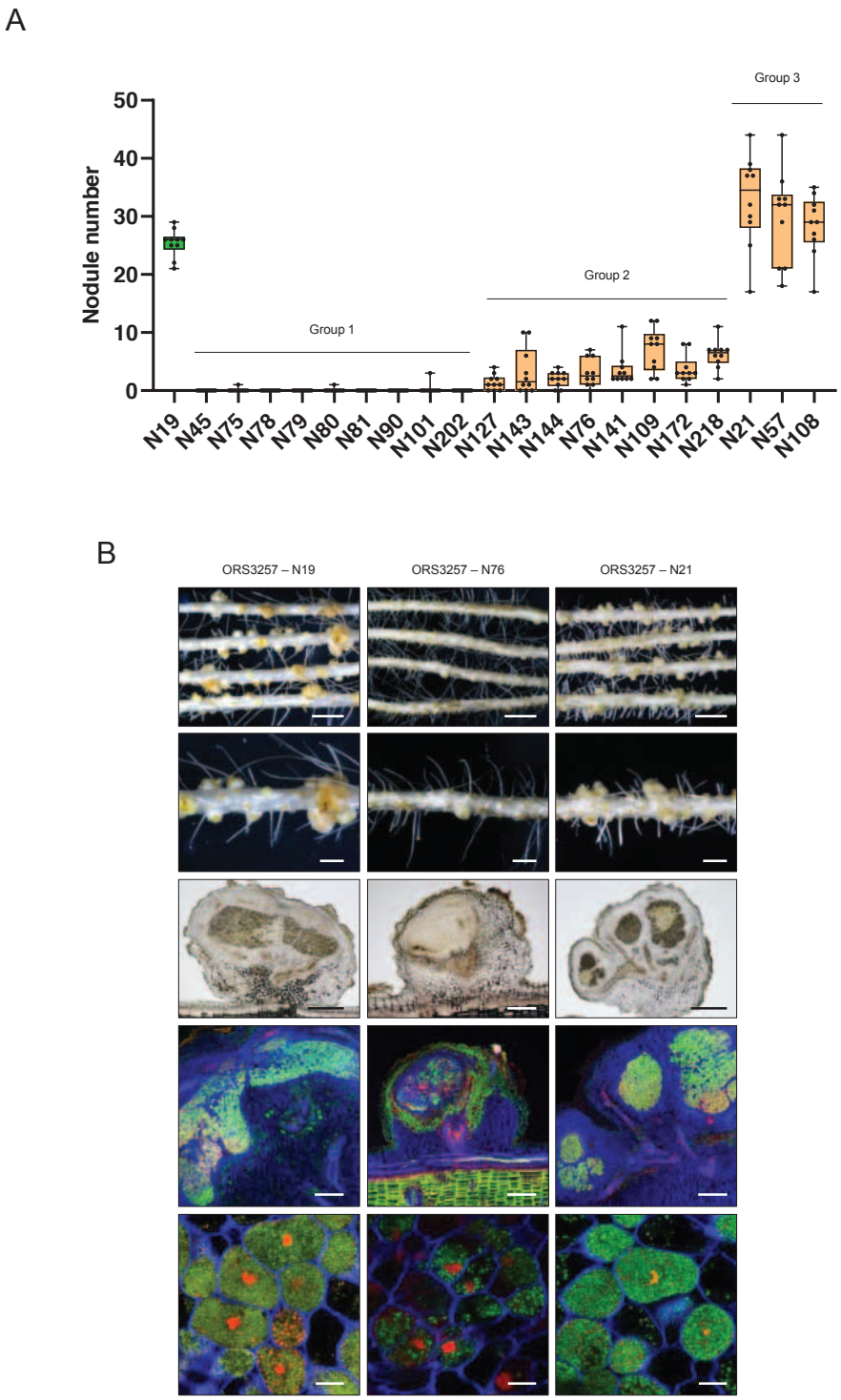


Figure 1. Identification of *Aeschynomene evenia* ecotype PI 225551 (N21) as an appropriate model for deciphering the T3SS dependent symbiotic process.

(**A**) Nodule number of different accessions of *A. evenia* ssp. *evenia* plants at 21 days after inoculation with the model strain ORS3257. Box plots show the results of one experiment performed on 10 plants for each accession. The green colour refers to *A. indica* N19 while the orange colour refers to *A. evenia* ssp. *evenia* accessions. (**B**) Symbiotic phenotypes of ORS3257 on 3 different *Aeschynomene* accession. N19: *A. indica* (LSTM 19), N76: *A. evenia* (CIAT 22838) and N21: *A. evenia* (PI 225551). In each case, the *Bradyrhizobium* strain is able to induce intracellularly infected nodules but the number of nodules is variable depending on the accession. Row 1 and 2: photo of roots and nodules. Scale bars: row 1, 0.5 cm; row 2, 0.2 cm. Row 3: Micro-sections of nodules observed using light microscopy. Scale bars: 250 µm. Row 4 and 5: confocal microscopy images of of nodule sections after staining with SYTO 9 (green, live bacteria), propidium iodide (red; infected plant nuclei and dead bacteria or bacteria with compromised membranes) and calcofluor (blue, plant cell wall). Scale bars: row 4, 200 µm; row 5, 10 µm.

reference genome *Aeschynomene_evenia*_v1.0 (Quilbe et al. 2021) using the NF-core/rnaseq pipeline (Ewels *et al*., 2020; Patel *et al*., 2020). Briefly, reads were aligned to the reference genome using STAR (Dobin *et al*., 2013) and gene expression was quantified with Salmon (Patro *et al*., 2017). Differentially expressed genes were identified using DESeq2 (Love et al., 2014) as implemented in the Diane pipeline (Cassan *et al*., 2021) and defined using a false discovery rate of 0.05 and fold change log2 ratio ≥1.

The Supplementary Table S3 gives for each library a summary of raw Illumina sequencing and filtered reads after trimming and alignment to *A. evenia* N76 genome.

Hairy root transformation with *Agrobacterium rhizogenes*

Plasmid *p35S-ernA* containing *ernA* under the control of the 35S promoter (Teulet et al. 2019), and the empty vector pJCV51 with the DsRed marker (https://gateway.psb.ugent.be), were used for hairy root transformations. *A. evenia* root transformation was performed following previously described procedures (Bonaldi et al. 2010b).

Results

***Aeschynomene evenia* ecotype PI 225551 (N21), an appropriate accession for deciphering the T3SS-symbiotic process.**

The model strain *B. vignae* ORS3257 was previously shown to nodulate in a NF/T3SS-dependent fashion *A. indica* LSTM19 (N19). This species represents a convenient host plant for nodulation tests but its allopoplyploid genome renders plant gene analysis intricated (Arrighi et al. 2012; Chaintreuil et al. 2018). To circumvent these complications, we aimed at identifying an accession in the closely related and diploid *A. evenia* species that equals the nodulation behaviour of *A. indica* N19. In a first approach, we examined a collection of 20 *A. evenia* accessions for their ability to be nodulated by strain *B. vignae* ORS3257. Three distinct groups can be distinguished (Fig. 1A): i) Group 1 with 9 accessions that are not or very rarely

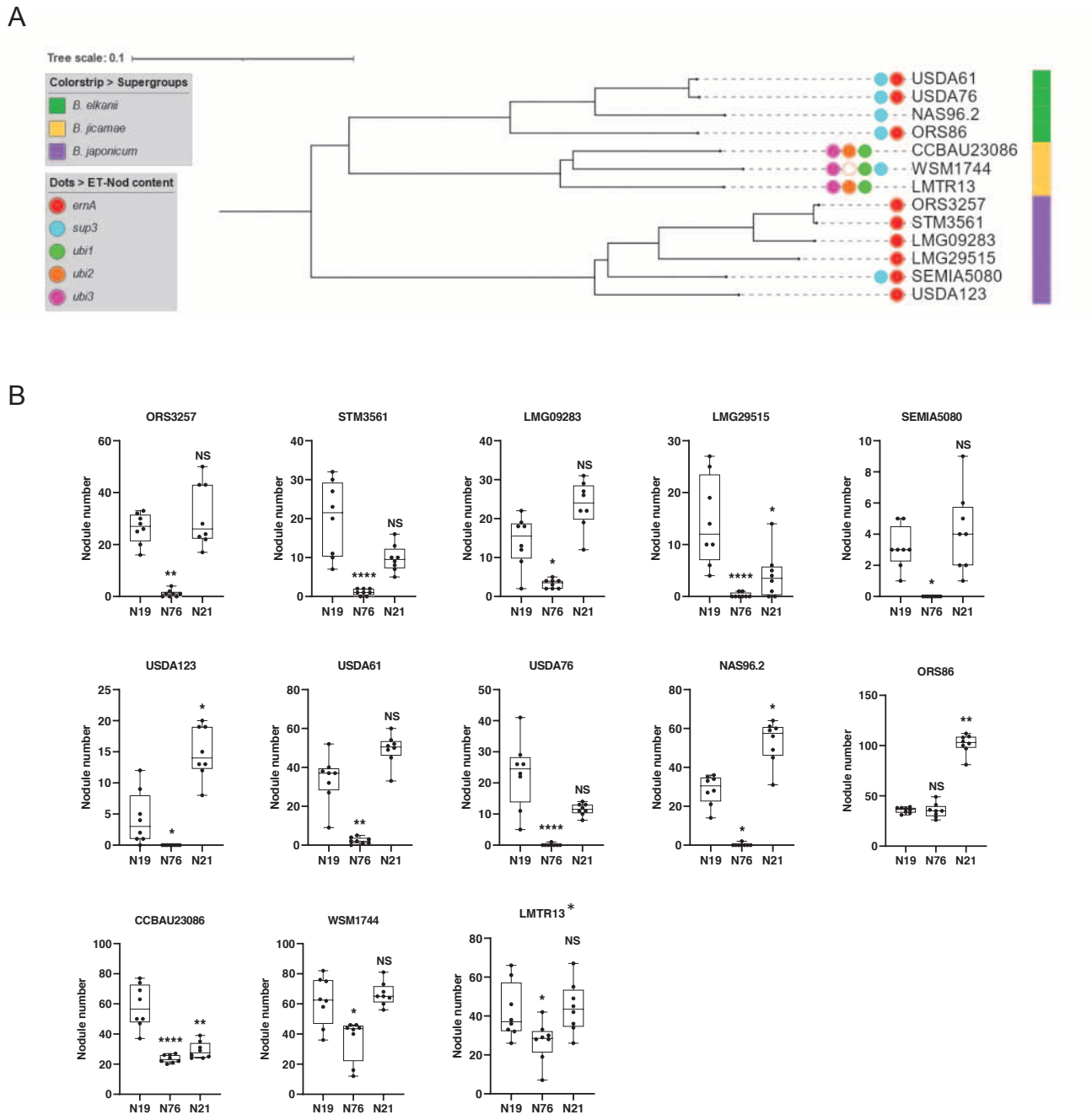


Figure 2. Symbiotic properties of various *Bradyrhizobium* strains inoculated on A. *indica* N19 and *A. evenia* N76 and N21.
(**A**) Phylogenomic tree of the 13 *Bradyrhizobium* strains used in this study obtained with the pruning tree option in iTOL (https://itol.embl.de/) (Letunic and Bork 2021), from a previously phylogenomic analysis (Camuel et al. 2023). Strains from three phylogenomic supergroups were used in our analysis: *B. japonicum* (6 strains), *B. elkanii* (4 strains) and *B. jicamae* (3 strains). The putative ET-Nods content for each strain is indicated by coloured dots referenced in the key (**B**) Nodule number induced by each strain on three different *Aeschynomene* accession plants (N19, N76 and N21) at 21 days after inoculation. Box plots show the results of one (representative) experiment performed on 8 plants for each accession. **** P<0.0001, ** P<0.01, * P<0.05, significant differences between the *A. indica* (N19) ecotype and the two others *A. evenia* accessions (N76 and N21) using a nonparametric Kruskal-Wallis test, NS: not significant. It is to note that the "pseudo-nodules" induced by LMTR13 on *A.evenia* N76 are extremely small and as a result the exact nodule number difficult to obtain.

nodulated, ii) Group 2 with 8 accessions, including N76, that are weakly nodulated with an average of 1 to 7 nodules per plant, and finally, iii) Group3 with 3 accessions (N21/N57/N108), that are well nodulated with an average of 25 to 30 nodules per plant, a level of nodulation comparable to the one observed on *A. indica* N19. Interestingly, these 3 last accessions are found grouped together in a NJ tree previously obtained and constructed from the allelic data of SSRs markers, which indicates their genetic proximity (Chaintreuil et al. 2018). While the geographical origin of one accession (N108) is unknown, the other two (N21 and N57) originate from the same region in West Africa (in the vicinity of the Victoria falls) (Table S1), suggesting that these three accessions may correspond to the same ecotype. As genomic sequence data are available for the N21 accession **(**Quilbé et al. 2022), we decided to focus on this one.

To characterize nodule formation with the strain ORS3257 in *A. evenia* N21, the nodule structure and bacterial occupancy was visualized by light and confocal microscopy, and compared to *A. evenia* N76 and to *A. indica* N19 (Fig. 1B). Nodule sections revealed that in each case, the nodules contained a well-defined central tissue with plant cells intracellularly infected with live bacteria (Fig. 1B).

Since the symbiotic phenotype (nodule number and the pattern of infection) strongly differs between *Bradyrhizobium* strains using the T3SS (Okazaki et al. 2016; Camuel et al. 2023), we compared the nodulation properties of 13 different *Bradyrhizobium* strains on two *A. evenia* accessions (N76 and N21), as well as on *A. indica* N19. The 13 strains were selected based on their phylogenetic position, their predictive ET-Nod content and their ability to nodulate *A. indica* N19 (Camuel et al. 2023) (Fig. 2A). As shown in Fig. 2B, for the majority of the strains belonging to the *B. japonicum* or *B. elkanii* phylogenetic supergroups, nodulation of *A. evenia* N76 was null or extremely reduced. Only the ORS86 strain and the 3 strains of the phylogenetic *B. jicamae* supergroup showed a significant level of nodulation (> 20 nodules per plant). Furthermore, these “nodules” remained often very small and their number was difficult to

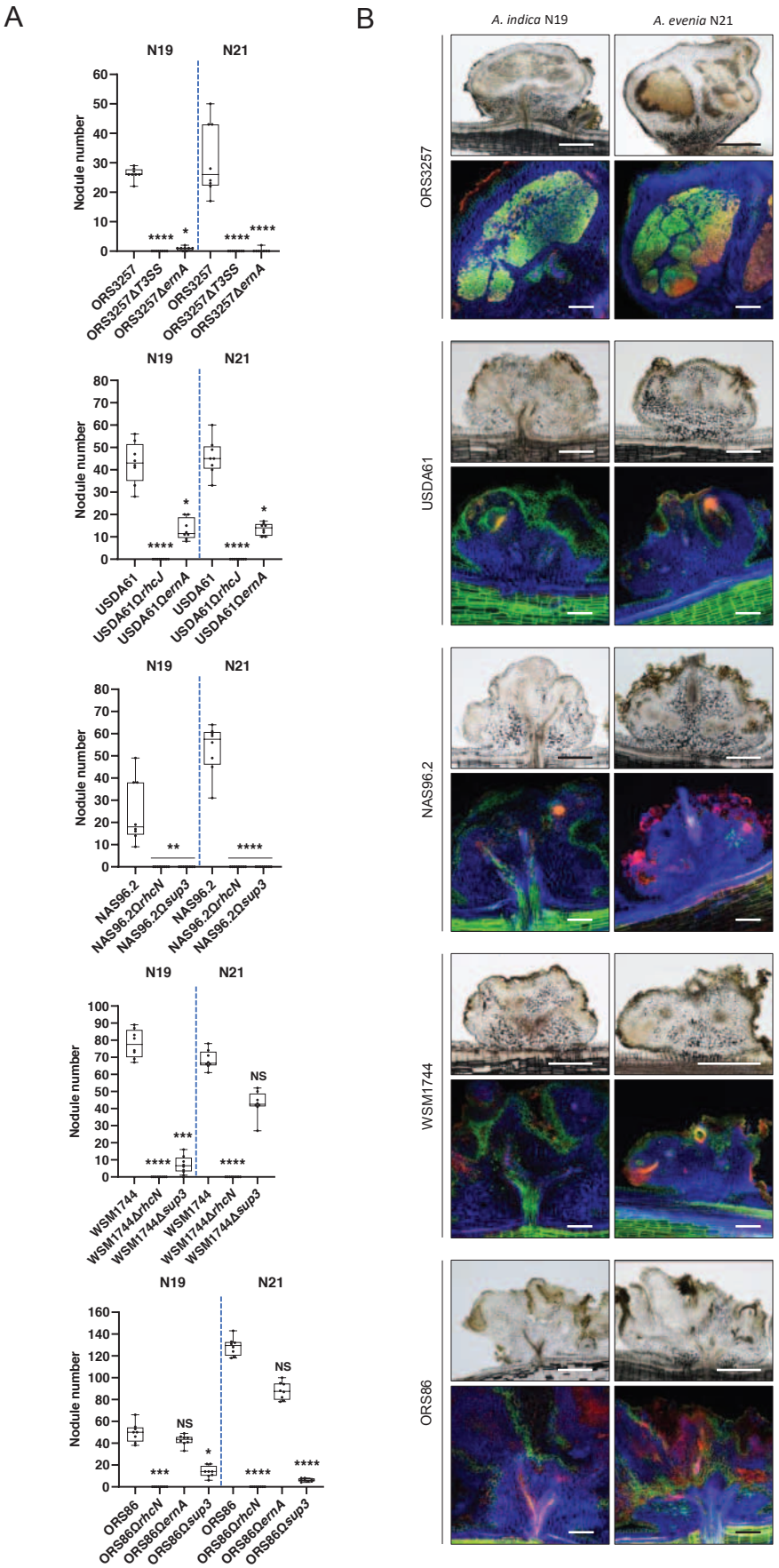


Figure 3. The symbiotic properties of various *Bradyrhizobium* strains and their respective mutants are conserved on *A. indica* N19 and *A. evenia* N21.

(**A**) Nodule number on *A. indica* N19 and *A. evenia* N21 plants at 21 days after inoculation with 5 *Bradyrhizobium* strains used in this study and their respective mutants. Box plots show the results of one (representative) experiment (8 plants each). * P<0.05, ** P<0.005, *** P<0.0001, **** P<0,0001, significant differences between the wild type strain and each mutant using a nonparametric Kruskal-Wallis test, NS: not significant. All the mutants tested were previously obtained (Okazaki et al. 2009; Teulet et al. 2019; Camuel et al. 2023). (**B**) Upper rows: Micro-sections of nodules observed using light microscopy. Scale bars: 500 µm. Lower rows: confocal microscopy images of nodule sections after staining with SYTO 9 (green, live bacteria), propidium iodide (red; infected plant nuclei and dead bacteria or bacteria with compromised membranes) and calcofluor (blue, plant cell wall). Scale bars: 200 µm.

determine with the naked eye. In contrast, on *A. evenia* N21, for the 13 strains tested, the number of nodules formed and their size was generally comparable to what observed on *A. indica* N19. To confirm that the nodulation observed on *A. evenia* N21 was T3SS-dependent and governed by the same ET-Nod as identified for *A. indica* N19, we analysed the symbiotic properties of various available mutant strains. As shown in Fig. 3, mutation of the T3SS (Ω*rhcN or* Ω*rhcJ* mutants) in 5 strains (ORS3257, USDA61, NAS96.2, WSM1744, ORS86) completely aborted the nodulation on *A. evenia* N21. Furthermore, strains with mutation in different ET-Nods have the same phenotype on *A. indica* N19 and *A. evenia* N21, except for the WSM1744 strain for which the *Sup3* mutation has a less drastic effect on *A. evenia* N21 (Fig. 3A). Of note, for the 5 strains of which mutants are tested, the pattern of nodule infection on *A. indica* N19 and *A. evenia* N21 is comparable (Fig. 3B).

For the development of certain functional approaches, such as ChIP-Seq, Immunoprecipitation or Turbo-ID proximity labelling to identify the nucleic or protein interactants of an ET-Nod, it may be interesting to use transgenic plants expressing ET-Nod in the absence of the bacteria and other interfering T3Es. We have previously shown that the ectopic expression of *ernA* in *A. indica* N19 roots activates organogenesis of root- and nodule-like structures. A similar response was observed in transgenic hairy root lines of *A. evenia* N21 expressing *ernA* from ORS3257 (Fig. 4). To put the boot in, the ectopic expression of the ET-Nod *sup3* from NAS96.2 strain was also previously shown to led also to the induction of pseudo-nodules on both *A. indica* N19 and *A. evenia* N21 (Camuel et al. 2023). Taken together these data indicate that *A. evenia* N21/PI 225551 is an appropriate ecotype to analyse the symbiotic properties of a large diversity of *Bradyrhizobium* strains using a NF/T3SS-dependent process and it can completely substitute *A. indica* N19, notably for the development of functional and transcriptomic approaches.

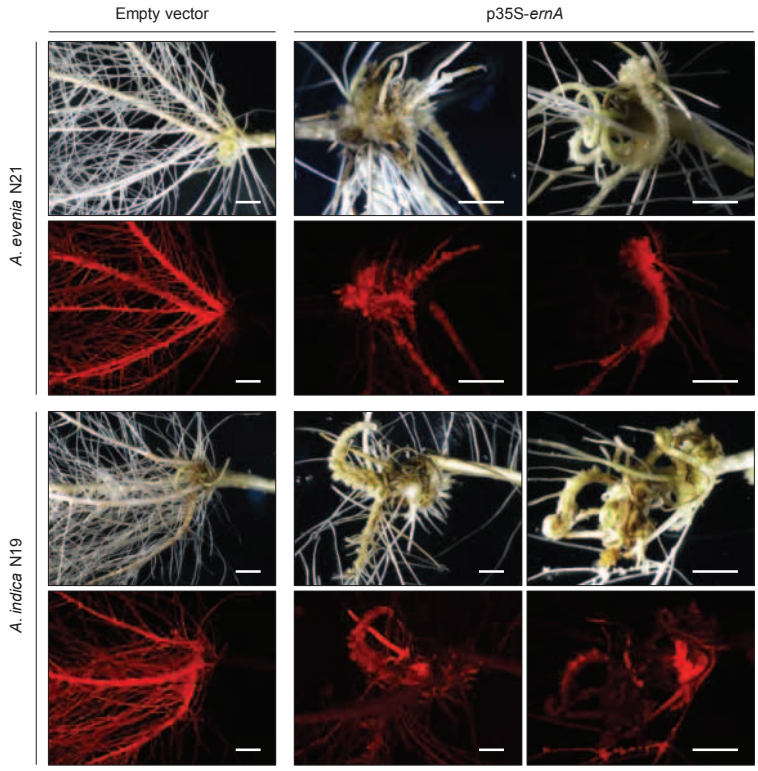
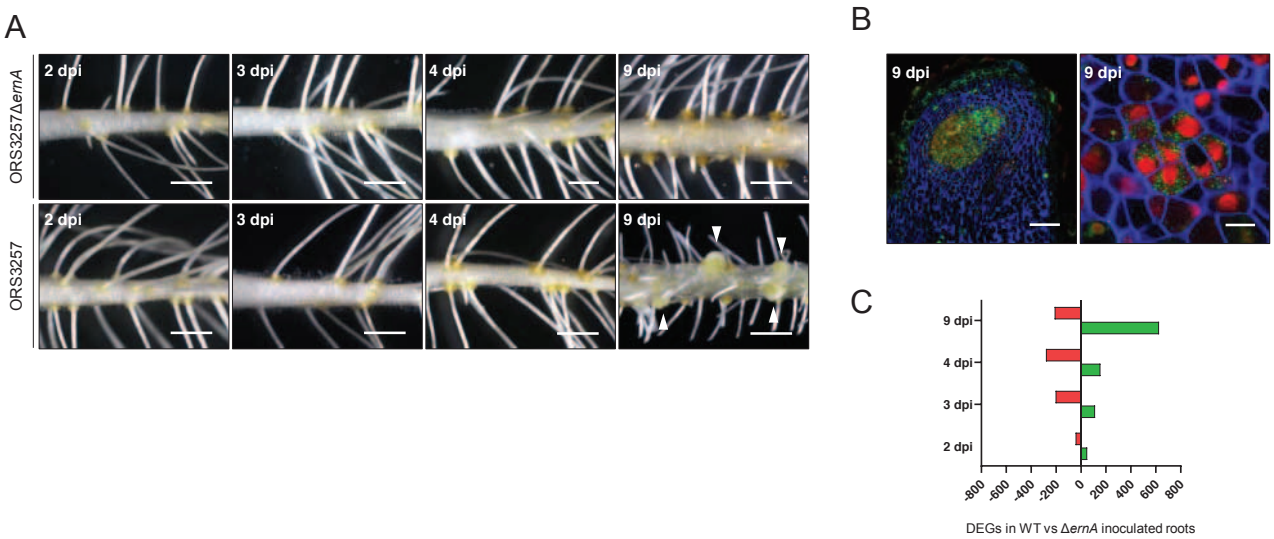


Figure 4. Ectopic expression of *ernA* from ORS3257 in *A. evenia* N21 and *A. indica* N19 hairy-roots.
Photos of *A. evenia* and *A. indica* roots transformed either with the empty vector containing the DsRed marker or with *p35S-ernA* at 6 weeks after transformation and in the absence of bradyrhizobia. Roots were observed by a fluorescence stereomicroscope equipped with a DsRed filter. Scale bars: 0.3 cm. In both cases, ErnA induces nodule-like structures as previously shown in *A. indica* (Teulet et al. 2019).

RNA-seq analysis reveal potential targets of ErnA

To better understand how ErnA activates nodule organogenesis, we compared the transcriptome of *A. evenia* N21 roots at different time points after inoculation with the WT-ORS3257 strain and the ORS3257Δ*ernA* mutant. We focused this analysis on very early stages of symbiosis (2, 3, 4 and 9 dpi). It is important to note that the NF/T3SS-dependent nodulation in *A. evenia* is a much slower process than the T3SS-independent process observed with photosynthetic *Bradyrhizobium* such as ORS278 or ORS285 strains. While nodules are visible from 3 dpi with the ORS285 strain and are perfectly functional as early as 5 dpi (Bonaldi et al. 2011), nodules are visible from 9 dpi with ORS3257 and no particular plant responses are visible by the naked eye at earlier time points (Fig. 5A).

Differentially expressed genes (DEG) between roots inoculated with the WT and ORS3257Δ*ernA* mutant at each time point were defined using a false discovery rate of 0.05 and fold change log2 ratio ≥1 (Fig. 5C and 5D and Table S4). While the number of up and down regulated DEGs (49/44) was comparable at 2 dpi, at 3 and 4 dpi there were about 2 times more down regulated genes than up regulated genes in the WT conditions as compared to the Δ*ern*A mutant conditions (Fig. 5C). At 9 dpi, this tendency is completely reversed with 626 up-regulated genes versus 212 down-regulated genes in WT inoculated roots compared to the ones inoculated with Δ*ernA* mutant. This suggests that in the very early stages, one of the primary functions of ErnA (direct or indirect) is to repress a number of biological processes. At 9 days post inoculation, the new wave of up-regulated genes observed in WT-inoculated roots is surely related to the fact that nodulation and infection were initiated at this time point (Fig. 5A and 5B). It can be assumed that the formation of the symbiotic organ and its bacterial infection results in a complete transcriptional reprogramming of the plant that is certainly partly beyond the control of ErnA .



D

	Categories	Gene ID	DEGs 2 dpi	3 dpi	4 dpi	9 dpi
Transcription factors	WRKY-TFs	Ae07g13440	6,4	-	-4,43	-
		Ae10g12720	1,43	-	-	1,46
		Ae01g30690	1,25	-	-	-
		Ae07g02180	-1	-1,1	-1,18	-
		Ae10g12700	-	-	1,22	2,22
		Ae08g05500	-	-	-1,25	-
		Ae08g04850	-	-	1,2	-
		Ae02g01930	-	-	-1,08	-
	Ethylene-responsive factors	Ae06g00820	2,25	2,21	1,73	2,29
		Ae09g17160	1,89	-	-	-
		Ae04g11400	1,27	-	1,28	-
		Ae06g00840	1,22	1,71	-	-
		Ae01g03740	-	1,28	-	-
		Ae06g29920	-	-2,34	-4,62	-
		Ae09g22590	-	-1,23	-	-
		Ae02g18900	-	-1,17	-1,92	-
		Ae07g18050	-	-1,08	-1,35	-1,05
		Ae04g37530	-	-	1,36	-
		Ae06g06120	-	-	-1,17	-
	Lateral root development	Ae08g03290	-	1,06	1,04	-
		Ae03g28880	-	1,81	1,17	-
	Others	Ae06g14010	1,79	1,41	-	1,28
		Ae08g04550	1,78	-	-	-
		Ae08g04260	1,22	-	-	-
		Ae06g29380	1,07	1,49	2,04	-
		Ae05g31640	-1,88	-2,74	-3,55	-1,01
		Ae10g11050	-1,84	-	-	-
		Ae08g02860	-1,28	-1,00	-	-
		Ae04g26630	1,24	-	-	-
		Ae04g26590	-	5,07	-	1,87
		Ae09g28280	-	4,96	5,5	3,1
		Ae03g14060	-	1,87	-	-
		Ae04g05640	-	1,74	-	1,40
		Ae01g02360	-	1,53	-	-
		Ae09g20520	-	1,48	1,01	-
		Ae02g03760	-	1,06	1,14	-
		Ae06g00890	-	-1,89	-	-
		Ae05g12640	-	-1,71	-2,13	-
		Ae07g14510	-	-1,68	-	-
		Ae03g18120	-	-1,43	-	-
		Ae08g11470	-	-1,25	-	-
		Ae06g01880	-	-1,23	-	-
		Ae04g12060	-	-1,18	-2,23	-
		Ae01g26370	-	-1,04	-1,01	-
		Ae01g18060	-	-	3,17	1,75
		Ae04g31250	-	-	1,95	1,02
		Ae02g22290	-	-	1,34	-
		Ae10g20290	-	-	-2,47	-
		Ae01g22710	-	-	-1,67	-
		Ae04g26600	-	-	-1,13	-
		Ae09g26250	-	-	1,1	-
Plant immunity	ACC-synthase	Ae06g32820		-10,73	-7,65	
	Peroxidases	Ae05g27780	-1,05	-	-	-
		Ae01g13530	-	-	-1,60	-
	TIR-NBS-LRR	Ae07g02640	-	-	-2,97	1,34
		Ae09g03250	-	-	-1,35	-
		Ae07g02630	-	-	-1,11	-
		Ae02g04130	-	-	-1,09	-
		Ae05g30930	-	-	-1,03	-
	Other R proteins	Ae01g29070	-	-1,72	-	-
		Ae04g36510	-	-1,63	-	-
		Ae01g00100	-	-1,02	-	-
Auxin-related genes	Putative auxin transporters	Ae01g25230	-1,36	-1,19	-1,77	-
		Ae07g18020	-	1,86	-	-
		Ae01g20380	-	1,33	-	-
		Ae04g13450	-	1,09	-	-
		Ae07g17980	-	-1,32	-1,00	-
		Ae03g22270	-	-1,28	-	-
		Ae07g18030	-	-1,05	-	-
		Ae01g25230	-1,36	-1,19	-1,77	-
	SAUR genes	Ae04g28180	-	-	-	2,22
		Ae04g34280	-	-	-	1,43
		Ae04g28160	-	-	-	1,19
		Ae04g34260	-	-	-	1,14
		Ae04g25150	-	-	1,97	-
		Ae06g04130	-	-	-1,55	-1,17
		Ae09g07430	-	-	-	-1,12
		Ae04g34370	-	-	-	-1,05

	Categories	Gene ID	DEGs 2 dpi	3 dpi	4 dpi	9 dpi
Signaling pathways	Receptor-like kinase	Ae01g03750	-	-	-1,54	-
		Ae09g24380	-	-	-1,46	-
		Ae08g25950	-1,66	-	-	-
		Ae05g08460	-	-2,85	-	-
		Ae02g20100	-	-2,37	-3,67	-
		Ae06g36630	-	-2,17	-	-
		Ae08g08160	-	-1,07	-1,30	-
		Ae09g10730	-	-	4,09	-
		Ae05g01000	-	-	1,23	-
		Ae09g24400	-	-	-1,40	-
		Ae09g24230	-	-	-1,09	-
		Ae01g00180	-	-1,03	-	-
		Ae09g22250	-	-	-1,43	1,01
		Ae08g08170	-	-	-1,31	-
		Ae09g25910	-	-	-1,07	-
		Ae04g06850	-	-	-1,04	-
		Ae05g26950	-	-	-1,03	-
		Ae06g36600	-	-	-1,02	-
		Ae01g23380	-	-	-1,02	-
		Ae03g02350	-	-	-1,40	-
	Protein kinase	Ae03g24610	-	-9,09	-	-8,72
		Ae0012210	-	-7,17	-	-4,91
		Ae09g05630	-	-2,74	-3,12	-
		Ae04g02860	-	-1,50	-2,93	-
		Ae06g30620	-	-1,07	-	-
		Ae09g25900	-	-	-2,13	-
		Ae04g12170	-	-2,67	-	-
		Ae07g06130	-	-2,37	-2,38	-
		Ae04g28970	-	-2,12	-2,70	1,79
		Ae02g20010	-	-	-1,43	-
		Ae05g34560	-	-	-1,20	-
		Ae09g25850	-	-	-1,13	-
		Ae09g17300	-	1,02	-	-
Symbiotic genes	AeLB1	Ae04g33090	-	-	-	4,14
	AeLB2	Ae04g33100	-	-	-	2,70
	AeNCR46	Ae06g25462	-	-	-	4,07
	AeNCR12	Ae02g17611	-	-	-	-2,02
	SYMREM1	Ae03g30480	-	-	-	3,14
	AeERN1	Ae06g15750	-	-	-	-1,07
	AeERN3	Ae03g25920	-	-	-	-1,34
	PUB1	Ae03g09100	-	-	-	-1,34
Others	Subtilases	Ae03g08090	8,46	-4,99	-6,43	-
		Ae04g28630	1,12	1,06	2,07	-
	Cell-wall modelling	Ae06g11170	-	-	1,67	3,52
		Ae08g26160	-	-	-	2,36
		Ae04g35860	-	-3,14	-	2,17
		Ae01g25730	-	-	-	1,70
		Ae10g08500	-	-	-	1,14

Figure 5. RNAseq analysis between *A. evenia* N21 roots inoculated with ORS3257 wild-type strain and its Δ*ernA* mutant.
(**A**) Symbiotic phenotype of root plants at 2, 3, 4 and 9 dpi inoculated with ORS3257 and ORS3257 Δ*ernA* mutant. Some nodules (indicated by white arrows) are visible on plant inoculated with WT strain at 9 dpi while no nodules are observed on plants inoculated with the Δ*ernA* mutant. Scale bars: 500 µm. (**B**) Images of micro-section of nodules elicited by the WT strain showing that nodules at 9 dpi were infected intracellularly. The observations were done using a confocal microscopy after staining with SYTO 9, propidium iodide and calcufluor (E-F). Scale bars: E, 200 µm; F, 10 µm. (**C**) Number of differentially expressed genes (DEG) in roots inoculated with the WT strain versus Δ*ernA* mutant at 2, 3, 4 and 9 dpi. The DEGs were defined using a false discovery rate of 0.5 and fold change log2 ratio ≥1. (**D**) Summary of the main categories of DEGs identified. A range colour from red to green indicates down-regulated to up-regulated genes in WT-inoculated roots Versus Δ*ernA*-inoculated roots.

At 2, 3 and 4 dpi, among the categories of genes found to be the most differentially expressed (up or down) between the WT inoculated roots and the Δ*ernA* mutant were transcription factors (TFs) with a high proportion of WRKY-TFs (8 genes) and ethylene-responsive factors (11 genes) (Fig. 5D and Table S4). These two families of TFs are known to be involved in plant immunity responses against pathogen attack (Tsuda and Somssich, 2015). In relation, the ACC synthase gene (Ae06g32820) which permits the synthesis of ethylene, the plant hormone defence, is one of the genes found the most strongly repressed in the WT condition. We also observe a down-regulation in WT roots of peroxidase encoding genes (Ae05g27780, Ae01g13530) that can be involved in reactive oxygen species (ROS) production leading to oxidative burst and cell death (Almagro et al. 2009). The repression of peroxidases could prevent a hypersensitive response (HR) that would be non-compatible with nodule organogenesis and infection. ErnA is therefore possibly involved in the repression of plant defence genes via the modulation of dedicated TFs in plant immunity control. It is also important to note the considerable number of disease resistance proteins (R proteins) that appear down-regulated at 3 and 4 dpi in WT inoculated roots. Among them, 5 TIR-NBS-LRR proteins have been identified. This class of R proteins are known to be involved in T3Es recognition and subsequent activation of effector triggering immunity (ETI) (Cui et al. 2015). In this sense, ErnA might prevent the recognition of others effectors injected in the host cell and avoid ETI.

The hormone auxin appears also be a key player during this interaction since the expression of several putative auxin transporters genes (8 genes) were found strongly modulated (some positively, others negatively) in the WT in comparison to the Δ*ernA* mutant conditions (Fig. 5D). Auxin influences nearly all aspects of plant growth and development including nodule organogenesis. The importance of auxin in nodulation has been well illustrated by the application of some auxin transporter inhibitors shown to induce early nodulines and nodule organogenesis (Hirsch et al., 1989; Rightmyer and Long, 2011). It is therefore possible that

ErnA acts by modulating the auxin content in roots through the modulation of expression of various auxin transporters. Noteworthy is also the up-regulation of several small auxin up-regulated RNA (SAUR) genes (8 genes) at 4 and 9 dpi in the WT conditions when nodulation is initiated. The SAUR genes constitute a family of auxin-responsive genes which corresponding proteins are considered as key effector outputs of auxin signal to regulate plant growth and development (Ren and Gray, 2015). Some of the identified SAUR genes could therefore be directly involved in triggering nodulation. Other TFs known to be specifically involved in lateral root development and auxin responses were found to be up-regulated in the WT-inoculated roots at 3 and 4 dpi, such as Ae08g03290 which encodes a putative MIZU-KUSSEI 1-like protein and Ae03g28880 which encodes for a putative SCARECROW-like protein (Moriwaki et al. 2011; Shaar-Moshe and Brady, 2023).

Among the other remarkable DEGs identified were genes involved in signaling pathways, including several putative receptor-like kinases (20 genes) and protein kinases (13 genes). Most of these genes were found downregulated at 3 and 4 dpi in the WT roots. The lower expression of these genes, which are most likely involved in the modulation of different signaling pathways in response to various input signals, could explain the general down-regulation of gene expression observed in WT- inoculated roots. ErnA could be directly involved in the control of some of these receptor kinases.

Among the genes that were found the most differentially expressed are also two genes encoding subtilisin-like serine proteases (Ae03g08090 and Ae04g28630). Subtilases are a large family of serine proteinases involved in various mechanisms in plants ranging from development to signaling during biotic interactions (Figueiredo et al., 2014; Schaller et al., 2017). Notably some subtilases were shown to be involved in some symbiotic processes, such as the subtilases SbtM1and SbtM3 in *Lotus japonicus* which play a role in the mycorrhiza infection process (Takeda et al. 2009) or the subtilases Ag12 and Cg12 of actinorhizal plants (*Alnus glutinosa*

and *Casuarina glauca*) that were found strongly expressed in the actinorhizal nodules during early infection stages (Ribeiro et al. 1995; Svistoonoff et al. 2003). A subtilase gene (CL433Contig1–Ae05g09230) was also found strongly expressed during the NF/T3SS-independent symbiosis between ORS278 strain and *Aeschynomene evenia* (Gully et al. 2018). It is therefore possible that the two subtilase genes Ae03g08090 and Ae04g28630 identified here are important in the NF/T3SS-dependent process.

The recent *A. evenia* genome sequencing has permitted to identify the orthologs of most known symbiotic genes characterized in the model legumes (*Medicago* and *Lotus*) (Quilbé et al 2021). We took benefit of this list of possible symbiotic genes to examine in more details their expression during this analysis. Only a few DEG emerged and most of them were found at 9 dpi when nodules started to emerge (Fig 5D). Among the up-regulated genes were found: i) two leghemoglobin encoding genes (AeLB1 and AeLB2), leghemoglobins are required to protect the nitrogenase complex from oxygen (Becana et al. 2020). It is worth noting that AeLB1 was the gene found the most upregulated at 9 dpi in the WT roots, ii) a gene encoding a Nodule cysteine rich peptide (AeNCR46), NCRs are involved in the differentiation process of bacteria into bacteroids, the active nitrogen-fixing form (Czernic et al. 2016) and iii) the remorin SYMREM1, shown to regulate the intracellular infection process of rhizobia (Lefebvre et al. 2010). Most probably also related to this infection step observed at 9 dpi, it is noted the up-regulation of several genes involved in cell-wall loosening (including a polygalacturonase, Ae08g26160; two glucan endo-beta-glucosidases, Ae04g35860 and Ae10g08500; a xyloglucan glucosylase, Ae01g25730; and a pectinesterase/pectinesterase inhibitor Ae06g11170), that most probably are required to facilitate the bacterial internalization (Fig. 4D).

A few other symbiotic genes but downregulated this time in WT-inoculated roots were also identified : i) a NCR gene (AeNCR12); ii) the two Ethylene transcriptional factors ERN1 and ERN3, described to play antagonist role in the expression control of NF-elicited genes (Cerri

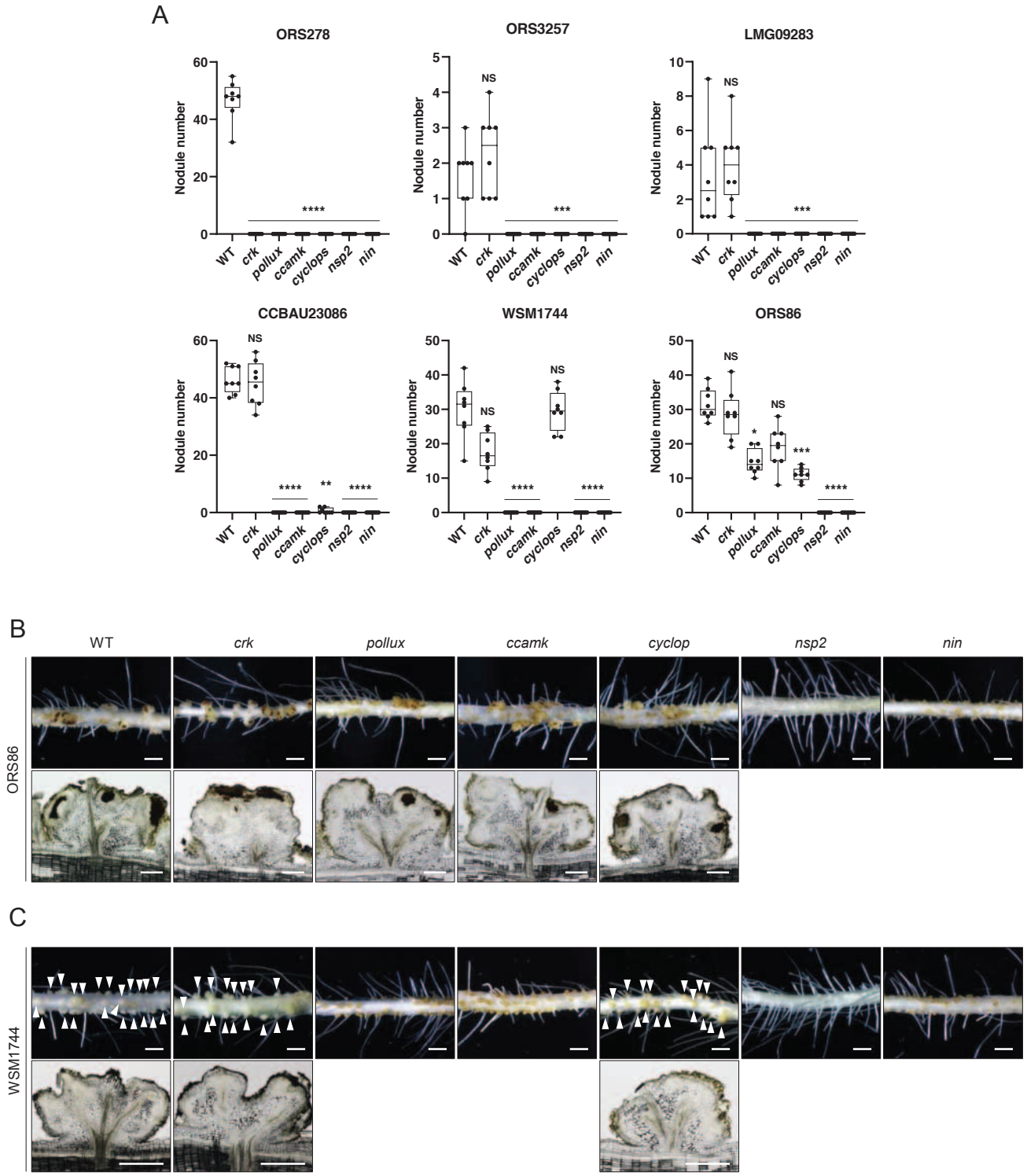


Figure 6. Some *Bradyrhizobium* strains can nodulate certain *A. evenia* N76 mutant lines affected in key symbiotic genes.
(**A**) Nodule number on wild-type *A. evenia* N76 and its different EMS mutant lines *(crk, pollux, ccamk, cyclops, nsp2, nin) after* inoculation with the photosynthetic *Bradyrhizobium* ORS278 strain using (T3SS-independent process) and 5 others *Bradyrhizobium* strains (ORS3257, LMG09283, CCBAU23086, WSM1744 and ORS86) able to nodulate *A. evenia* N76 using their T3SS. Observations were made at 21 days post-inoculation. Box plots show the results of one experiment (8 plants each). * P≤0.01, ** P≤0.001, *** P≤0.0005, **** P<0,0001, significant differences between the wild type strain and each mutant using a nonparametric Kruskal-Wallis test, NS: not significant. (**B**) Row 1: photo of roots and nodules induced by ORS86 on *A. evenia* N76 plants and its derivated EMS mutants. Row 2: Micro-sections of nodules observed using light microscopy. Scale bars: row 1, 0.2 cm ; row 2, 250 µm. (**C**) Row 1: photo of roots and nodules induced by WSM1744 on *A. evenia* N76 plants and its derivated EMS mutants. Small nodules are indicated by white arrows. Row 2: Micro-sections of nodules observed using light microscopy. Scale bars: row 1, 0.2 cm ; row 2, 250 µm.

et al. 2012) and iii) the E3 ubiquitin ligase-PUB1 that negatively regulates rhizobial and mycorrhizal symbioses in *Medicago truncatula* (Vernié et al. 2016).

Above transcriptomic data provide a general overview of the functions that could be modulated by ErnA and provide a first set of candidate genes whose expression could be directly targeted by it.

Some *Bradyrhizobium* strains thanks to their T3SS can nodulate *A. evenia* N76 lines mutated in *CRK*/*POLLUX*/*CCaMK*/*CYCLOPS* but not in *NSP2* and *NIN*.

Although *A. evenia* N76 is poorly nodulated by bradyrhizobia using their T3SS, the availability of EMS mutant lines in various symbiotic genes that are required for the NF/T3SS-independent process makes this accession interesting to investigate whether these determinants are also implicated in the T3SS-dependent nodulation process.

We analysed 5 *Bradyrhizobium* strains that previously have been shown to induce nodules on the WT *A. evenia* N76, on 6 mutant lines (*crk, pollux, ccamk, cyclops, nsp2* and *nin*). For the 5 analyzed strains, a number of common observations could be done (Fig. 6A): i) no significant difference in the number of nodules was observed between the WT and the *crk* mutant and ii) no nodulation was observed for the *nsp2* and *nin* mutants.

These data are very instructive as they indicate on the one hand that CRK is dispensable in the case of the T3SS-dependent symbiotic process and, on the other hand, that this signaling pathway requires the NIN and NSP2 TFs like the other symbiotic nitrogen-fixing processes described to date.

The impact of the mutations in the symbiotic determinants POLLUX, CCaMK and CYCLOPS depended on the strain tested: i) for the 2 strains (ORS3257 and LMG09283), no nodules were observed on these 3 mutant lines, ii) for the 2 strains (CCBAU23086 and WSM1744), the *pollux* and *ccamk* mutations completely aborted nodulation, on the other hand a few nodules could still be observed in the *cyclops* mutant for CCBAU23086, and this mutation has no effect on

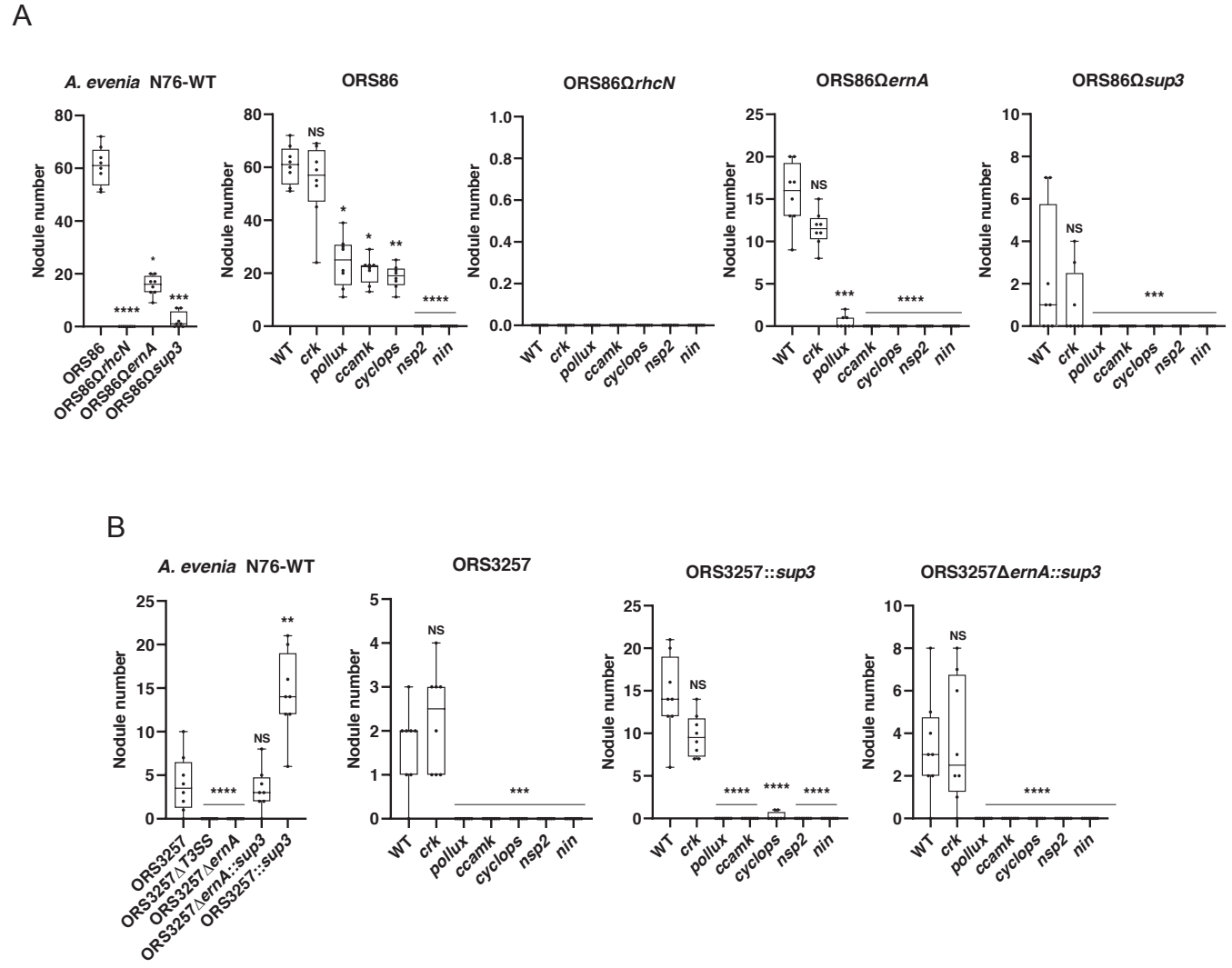


Figure 7. Synergetic role of ET-Nods during T3SS-dependent symbiosis.
(**A**) Nodule number on wild-type *A. evenia* N76 and its different EMS mutant lines *(crk, pollux, ccamk, cyclops, nsp2, nin*) after inoculation with ORS86 and its respective mutants in T3SS (Ω*rhcN*) and the two identified ET-Nods (Ω*ernA*, Ω*sup3*). Box plots show the results of one experiment (8 plants each). * P<0.05, *** P<0.001, **** P<0,0001, significant differences between the wild type strain and each mutant using a nonparametric Kruskal-Wallis test, NS: not significant. (**B**) Effect of the addition of the *sup3* gene of WSM1744 into ORS3257 and its Δ*ernA* mutant on the nodulation of wild-type *A. evenia* N76 and its different EMS mutant lines. Box plots show the results of one representative experiment (8 plants each). **P<0.01, *** P<0.001, **** P<0,0001, significant differences between the wild type strain or the wild type plant and each mutant using a nonparametric Kruskal-Wallis test, NS: not significant.

nodulation for WSM1744 (Fig. 6A and 6B), and finally iii) for the ORS86 strain, nodules could be observed on all 3 mutant lines but with a significantly lower number than the wild type (Fig. 6A and 6C).

These results indicate that some canonical symbiotic determinants are dispensable in some strains to induce nodule organogenesis in *A. evenia*.

Cooperativity between ET-Nods for nodulation activation

These observed differences between the strains can be explained by the fact that each of them can inject a specific set of T3Es. In particular, the ORS86 strain, which can bypass the mutation of several canonical symbiotic genes, has two ET-Nods (ErnA and Sup3). We examined in this strain the effect of the mutation of each ET-Nod as well as the effect of the T3SS mutation. As expected, inactivation the T3SS completely abolished nodulation, confirming that the nodules induced on wild type *A. evenia* N76 plants are dependent on T3SS (Fig. 7A). Inactivation of *ernA* or *sup3* significantly reduced the number of nodules on WT N76 plants suggesting that these two ET-Nods act synergistically to activate nodulation. This synergistic effect seems to be essential to bypass the mutation of plants in *POLLUX, CCaMK, CYCLOPS*. Indeed, the Ω*sup3* mutant was unable to induce nodules in these three mutant lines while the Ω*ernA* mutant only induced a few nodules in the *pollux* mutant (Fig. 7A).

We also examined the effect of adding the $sup3_{WSM1744}$ gene into strain ORS3257 and its ORS3257Δ*ernA* derivative mutant. As shown in Fig. 6B, addition of *sup3* in the ORS3257 WT strain resulted in a 4 times higher nodulation capacity on WT N76 plants. Nevertheless, on the plant mutant lines tested, no nodules were formed, except for *crk* mutant for which a nodulation similar to the WT line was observed, and for the *cyclops* mutant for which occasionally a nodule could be observed on some plants (2 plants on 8 with one nodule). In the ORS3257Δ*ernA* mutant, the addition of $sup3_{WSM1744}$ restored nodulation to a level comparable to the WT strain but did not facilitate nodulation of the *cyclops* mutant (Fig. 7B).

Taken together, these data indicate that ET-Nods can have a cumulative effect on activation of nodule organogenesis and allows some strains, to nodulate (more or less easily) plants with mutations in key symbiotic determinants such as POLLUX, CCaMK and CYCLOPS.

Discussion

Identification of *A. evenia* N21 ecotype as model plant to study the T3SS-dependent symbiotic process

The discovery of a novel alternative symbiotic pathway between bradyrhizobia and legumes based on the T3SS is recent (Okazaki et al. 2013). If progresses have been made these last years on the bacterial partner through the identification of various T3Es triggering nodulation in the absence of NFs, nothing is known on their mode of action and on the plant pathways activated after their translocation into the host cell.

So far, only certain species of two legume genera have been shown to have the capacity to be nodulated via a T3SS-dependent process : i) the *Glycine max* (soybean) cultivar Enrei which can be nodulated by a *nodC* mutant of *B. elkanii* USDA61 (Okazaki et al. 2013) and ii) several *Aeschynomene* species including *A. indica* which has been preferentially used given its robust nodulation response to a large diversity of *Bradyrhizobium* strains (Okazaki et al. 2016).

Like most legumes, *Glycine max* uses in nature a NF-dependent process. Thus, to decipher the molecular mechanism specific to the T3SS process in this legume species, the use of a plant mutant in the NF receptors (like *nfr1* (Ikeda et al. 2008)), and/or a bacterial mutant affected in the *nod* genes is necessary (Okazaki et al. 2013). In contrast, in the NF-independent *Aeschynomene* species an interference with activation of the NF-dependent signaling pathway is less likely. Although genes encoding NF receptors (NFP/LYK3) are present in *A. evenia* and related species, several indications suggest that these genes are not functional (Quilbé et al. 2022). Furthermore, *nod* gene mutants of strains USDA61 and ORS285 nodulate *A. indica* as

well as the original wild type strains showing that NFs are not necessary to activate nodulation in such species (Bonaldi et al. 2012; Okazaki et al. 2016). This is consistent with the fact that NF-independent *Aeschynomene* spp. are naturally nodulated by photosynthetic bradyrhizobia lacking *nod* genes (Giraud et al. 2007).

Among the NF-independent *Aeschynomene* species, *A. evenia* was chosen as model plant to decipher at the plant level the molecular mechanism of the NF/T3SS-independent process used by photosynthetic bradyrhizobia (Arrighi et al. 2012). Unfortunately, the selected *A. evenia* accession (N76) is poorly nodulated by the majority of *Bradyrhizobium* strains nodulating *A. indica* using their T3SS (Okazaki et al. 2016). This bradyrhizobial selectivity is shared by most of the accessions tested here except for 3 accessions, including N21, which most likely belong to the same ecotype. Strikingly, *A. evenia* N21, like *A. indica*, showed a high nodulation efficiency and was permissive to a large set of *Bradyrhizobium* strains using their T3SS. What determinant(s) control(s) the promiscuity of this ecotype would be an interesting question to investigate in the future as it could provide important new information on the T3SS-dependent process.

RNAseq analysis reveals a dual role of ErnA in plant defense responses and nodule organogenesis.

To progress in the understanding of the genes modulated by ErnA, we compared the transcriptome of *A. evenia* N21 roots inoculated with the ORS3257 WT or its derivative Δ*ernA* mutant. Two main functional groups emerged from this analysis: the repression in the roots inoculated by the WT-strain of numerous plant defense genes and the modulation, positively or negatively, of several genes related in auxin transport and signaling, a key hormone involved in root and nodule development (Mathesius 2008). Interestingly, a similar transcriptomic analysis was recently performed in *Glycine max nfr1* roots inoculated with USDA61 WT-strain and its derivative *bel2-5* mutant, Bel2-5 was shown to be the key T3Es governing nodulation

during this last interaction (Ratu et al. 2022). This study also revealed the repression of plant defense genes in WT-inoculated *Glycine max* roots. While auxin-related genes were not shown to be differentially expressed, several genes involved in cytokinin biosynthesis are up-regulated in the WT conditions. As the hormone cytokinin also plays a key role in nodule organogenesis (Boivin et al. 2016), we hypothesize that the ET-Nods ErnA and Bel2-5 play a similar dual role. On one hand they repress the plant immune system to allow intracellular acceptance of bacteria, and on the other hand they activate nodule organogenesis by modulating the cytokine/auxin balance but using different pathways.

This RNAseq analysis highlights a long list of candidate genes that could be direct targets of ErnA, including genes encoding transcription factors, genes involved in plant signaling or genes related to plant hormone synthesis or signaling. One of the precautions to be taken in this type of experiment is that the bacterium injects a cocktail of effectors into the host cell and that these effectors may interact antagonistically or synergistically (Teulet et al. 2022). Although an isogenic Δ*ernA* mutant is used, it cannot be excluded that the observed effects are the consequence of other effectors that arise due to the absence of ErnA. To overcome this type of "artefact" and to narrow the list of possible ErnA targets, experiments are in progress to conduct a transcriptomic analyses of *A. evenia* roots overexpressing ErnA using a dexamethasone-inducible promoter (pDEX). The fact that roots of the *A. evenia* N21 ecotype overexpressing ErnA, like *A. indica* roots, form nodule-like structures is encouraging to pursue this type of approach (Fig. 4).

Some key determinants of the NF-signaling pathway are recruited during the T3SS-dependent process

To further understand this alternative T3SS-dependent symbiotic process, we took advantage of several *Aeschynomene evenia* N76 mutant lines unable to form nodules with the model strain ORS278 that uses a NF / T3SS-independent process. Analysis of the nodulation phenotype of

theses mutants after inoculation with several *Bradyrhizobium* strains using a T3SS-dependent process and containing different ET-Nods for nodule induction provides several general insights. Firstly, the *AeCRK* determinant encoding a cysteine-rich receptor kinase (CRK), which has been shown to be necessary for NF-independent nodulation by strain ORS278 (Quilbé et al. 2021 and 2022), is dispensable in the case of strains using a T3SS-dependent process. CRK is thought to be part of a receptor complex sensing non-Nod symbiotic signal(s) synthesized by photosynthetic bradyrhizobia that are still remain to be disclosed. Bradyrhizobia that use a T3SS-dependent process and inject a set of effectors directly into the host cell thus probably do not need this initial recognition step by plant membrane receptors to activate nodulation. These observations reinforce the idea that *Ae*CRK is a specific key determinant of NF/T3SS-independent symbiosis involved in the very early stages of this symbiosis.

Secondly, the NSP2 and NIN transcriptional factors are key determinants of the T3SS-dependent pathway since all the 5 tested strains were Nod$^-$ on the corresponding mutants. These data suggest that the different ET-Nods identified in these 5 strains (ErnA, Sup3, Ubi1, Ubi2 and Ubi3) all act, in combination or alone, upstream of these two TFs to activate the nodulation process.

Thirdly, some canonical symbiotic genes are dispensable for some *Bradyrhizobium* strains to activate nodulation in a T3SS dependent manner. For example, strain WSM1744, and strain CCBAU23086 at a lower level, can induce nodules on a *cyclops* mutant. This observation is surprising as these 2 strains have a Nod$^-$ phenotype on the *ccamk* mutant while in the NF signaling pathway CYCLOPS acts downstream of CCaMK. In this pathway, CYCLOPS after phosphorylation by CCaMK activates the expression of related symbiotic genes including NIN (Singh et al. 2014). To explain our data, we could envisage that CCaMK phosphorylates other actors than CYCLOPS that also play a prominent role in the activation of nodulation. The hypothesis that CCaMK has other yet unknown targets (X) has already been proposed by

Limpens and Bisseling (2014) to explain the divergent symbiotic phenotypes observed between *cyclops* and *ccamk* mutants in *Lotus japonicus*.

The symbiotic phenotype of strain ORS86 is also amazing as this strain is able to by-pass the requirement of several key symbiotic genes (POLLUX, CCaMK, CYCLOPS) for nodulation. This strain has previously been shown to be one of the most efficient strains in nodule induction on *A. indica* plants and in the same was found with the *A. evenia* N76 line. The ORS86 strain has two ET-Nods (ErnA and Sup3) and we showed that both of them are required to bypass the POLLUX, CCaMK, and CYCLOPS mutations, suggesting that they act synergistically to trigger nodulation. At this stage, it is difficult to envisage how this cooperativity between these two ET-Nods works. Do they act at different levels of the signaling pathway or do they intervene on the same actor(s) and by cumulating their action allows them to reach a sufficient "threshold" level to trigger the nodule organogenesis process? These important questions remain to explored in future research to better understand the T3SS-dependent process.

In conclusion, this study identified the *A. evenia* N21 accession as model plant for functional analyses to elucidate the molecular mechanisms of the T3SS-dependent nodule organogenesis. While this research can benefit of all the genetic tools and data acquired on the closely related *A. evenia* N76 that is used to study the NF/T3SS-independent nodulation process, the promiscuous nodulation character of the N21 accession makes it highly fitting to investigate the multiple mechanisms by which different ET-Nods can induce nodulation.

The identification of possible targets of ErnA by the RNAseq analysis as well as the evidences that T3SS-dependent nodule organogenesis requires some classical symbiotic determinants such as NSP2 and NIN and that others can be dispensable constitute major advances in the understanding of this process. Identifying the targets of ET-Nods and understanding their mode of action are important knowledge gaps that need to be filled now. This knowledge will provide insights into the functioning of the root nodule symbiosis and could contribute to the

development of a more sustainable agriculture. Furthermore, understanding by which mechanisms the bacterium can bypass several key players in the symbiotic/nodulation signaling pathway while continuing to induce nodule organogenesis may also be an important avenue to explore in the future to transfer the symbiotic capacity to cereals.

Data Availability

The data of the RNAseq analysis are available at NCBI BioProject (accession: XXXXXX).

Acknowledgements

This work was supported by a grant from the French National Research Agency ("ET-Nod"; ANR-20-CE20-0012). A.C. was supported by a PhD fellowship from the French Ministry of National Education, Higher Education and Research.

References

- L. Almagro, L. V. Gómez Ros, S. Belchi-Navarro, R. Bru, A. Ros Barceló, M. A. Pedreño. 2009. Class III peroxidases in plant defence reactions. Journal of Experimental Botany, Volume 60, Issue 2, February 2009, Pages 377–39
- Arrighi JF, Cartieaux F, Brown SC, Rodier-Goud M, Boursot M, et al. (2012) Aeschynomene evenia, a model plant for studying the molecular genetics of the Nod-independent rhizobium-legume symbiosis. Mol Plant Microb Interac 25: 851– 61.
- Becana, M., Yruela, I., Sarath, G., Catalán, P. & Hargrove, M.S. Plant hemoglobins: a journey from unicellular green algae to vascular plants.New Phytol. 227, 1618–1635 (2020).
- Boivin S, Fonouni-Farde C, Frugier F (2016) How auxin and cytokinin phytohormones modulate root microbe interactions. Front Plant Sci 7:1240. doi: 10.3389/fpls.2016.01240.
- Bonaldi K, Gourion B, Fardoux J, Hannibal L, Cartieaux F, Boursot M, et al. (2010a) 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. 2010; 23:760–770.

- Bonaldi K, Gherbi H, Franche C, Bastien G, Fardoux J, Barker D, Giraud E, Cartieaux F (2010b) The Nod factor-independent symbiotic signaling pathway: development of *Agrobacterium rhizogenes*-mediated trans- formation for the legume *Aeschynomene indica*.Mol Plant Microbe In- teract 23: 1537–1544
- Bonaldi K, Gargani D, Prin Y, Fardoux J, Gully D. 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(11):1359-1371. doi: 10.1094/MPMI-04-11-0093.
- Camuel A, Teulet A, Carcagno M, Haq F, Pacquit V, Gully D, Pervent M, Chaintreuil C, Fardoux J, Horta-Araujo N, Okazaki S, Ratu STN, Gueye F, Zilli J, Nouwen N, Arrighi JF, Luo H, Mergaert P, Deslandes L, Giraud E. 2023. Widespread *Bradyrhizobium* distribution of diverse Type III effectors that trigger legume nodulation in the absence of Nod factor. ISME J. 2023 Jun 24. doi: 10.1038/s41396-023-01458-1.
- Cassan, O., Lèbre, S. & Martin, A. Inferring and analyzing gene regulatory networks from multi-factorial expression data: a complete and interactive suite. BMC Genomics 22, 387 (2021). https://doi.org/10.1186/s12864-021-07659-2
- Cerri MR, Frances L, Laloum T, Auriac MC, Niebel A, Oldroyd GE, Barker DG, Fournier J, de Carvalho-Niebel F (2012) *Medicago truncatula* ERN transcription factors: regulatory interplay with NSP1/NSP2 GRAS factors and expression dynamics throughout rhi- zobial infection. Plant Physiol 160: 2155–2172
- Chaintreuil C, Perrier X, Martin G, Fardoux J, Lewis GP, Brottier L, Rivallan R, Gomez-Pacheco M, Bourges M, Lamy L, Thibaud B, Ramanankierana H, Randriambanona H, Vandrot H, Mournet P, Giraud E, Arrighi JF. Naturally occurring variations in the nod-

independent model legume *Aeschynomene evenia* and relatives: a resource for nodulation genetics. BMC Plant Biol. 2018 Apr 3;18(1):54.

- Chomczynski, Piotr, et Nicoletta Sacchi. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction . Analytical Biochemistry 162, no 1 (1987): 156-59. https://doi.org/10.1016/0003-2697(87)90021-2.
- Cui H, Tsuda K, Parker JE. 2015. Effector-triggered immunity: from pathogen perception to robust defense. Annu. Rev. Plant Biol. 66:487-511. doi: 10.1146/annurev-arplant-050213-040012
- Czernic, P. 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.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner Bioinformatics. 2013 Jan 1;29(1):15-21. doi: 10.1093/bioinformatics/bts635. Epub 2012 Oct 25. PubMed PMID: 23104886; PubMed Central PMCID: PMC3530905.
- Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, Garcia MU, Di Tommaso P, Nahnsen S. The nf-core framework for community-curated bioinformatics pipelines. Nat Biotechnol. 2020 Mar;38(3):276-278.
- Feng J, Lee T, Schiessl K, Oldroyd GED.2021. Processing of NODULE INCEPTION controls the transition to nitrogen fixation in root nodules. Science.;374(6567):629-632.
- Figueiredo A, Monteiro F, Sebastiana M. (2014). Subtilisin-like proteases in plant-pathogen recognition and immune priming: a perspective. Front Plant Sci. 2014 Dec 19;5:739. doi: 10.3389/fpls.2014.00739. eCollection 2014.
- Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E, Avarre JC et al. Legumes symbioses: absence of *nod* genes in photosynthetic bradyrhizobia. Science. 2007; 316:1307-1312.

- Gully D, Czernic P, Cruveiller S, Mahé F, Longin C, Vallenet D, François P, Nidelet S, Rialle S, Giraud E, Arrighi JF, DasGupta M, Cartieaux F. 2018. Transcriptome profiles of Nod factor-independent symbiosis in the tropical legume *Aeschynomene evenia*. Sci Rep. 2018 Jul 19;8(1):10934. doi: 10.1038/s41598-018-29301-0.
- Hirsch AM, Bhuvaneswari TV, Torrey JG, Bisseling T (1989) Early nodulin genes are induced in alfalfa root outgrowths elicited by auxin transport inhibitors. Proc Natl Acad Sci USA 86: 1244–1248
- Ikeda S, et al. (2008) Microbial community analysis of field-grown soybeans with different nodulation phenotypes. Appl Environ Microbiol 74(18):5704–5709.
- Lefebvre B, Timmers T, Mbengue M, Moreau S, Herve C, Toth K, Bittencourt-Silvestre J, Klaus D, Deslandes L, Godiard L, et al. (2010) A remorin protein interacts with symbiotic receptors and regulates bacterial infection. Proc Natl Acad Sci U S A 107: 2343–2348
- Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021; 49: W293-W296.
- Limpens E, Bisseling T. (2014) CYCLOPS: a new vision on rhizobium-induced nodule organogenesis. Cell Host Microbe. 15(2):127-9.
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. Genome Biol. 15, 1–21 (2014).
- Mathesius U. (2008) Auxin: at the root of nodule development? Funct Plant Biol.;35(8):651-668.
- Moriwaki T, Miyazawa Y, Kobayashi A, Uchida M, Watanabe C, Fujii N, Takahashi H. 2011. Hormonal regulation of lateral root development in Arabidopsis modulated by MIZ1 and requirement of GNOM activity for MIZ1 function. Plant Physiol. 2011 Nov;157(3):1209-20. doi: 10.1104/pp.111.186270. Epub 2011 Sep 22.

- Oldroyd GED. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol. 2013; 11:252-263.
- Okazaki S, Kaneko T, Sato S, Saeki K. Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. Proc Natl Acad Sci USA. 2013; 110:17131-17136.
- Okazaki S, TIttabutr P, Teulet A, Thouin J, Fardoux J, Chaintreuil C. et al. Rhizobium–legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. ISME J. 2016; 10:64-74.
- Patel H, Ewels P, Peltzer A, Hammarén R, et al. nf-core/rnaseq: nf-core/rnaseq v3.0—SilverShark https://doi.org/10.5281/zenodo.4323183 (2020).
- Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression Nat Methods. 2017 Apr;14(4):417-419. doi: 10.1038/nmeth.4197. Epub 2017 Mar 6. PubMed PMID: 28263959; PubMed Central PMCID: PMC5600148
- Perret X, Staehelin C, Broughton WJ. 2000. Molecular basis of symbiotic promiscuity. *Microbiol. Mol. Biol.* 64(1):180-201. doi: 10.1128/mmbr.64.1.180-201.2000
- Quilbé J, Lamy L, Brottier L, Leleux P, Fardoux J, Rivallan R, Benichou T et al. . Genetics of nodulation in *Aeschynomene evenia* uncovers mechanisms of the rhizobium-legume symbiosis. Nat Commun. 2021 Feb 5;12(1):829. doi: 10.1038/s41467-021-21094-7.
- Quilbé J, Nouwen N, Pervent M, Guyonnet R, Cullimore J, Gressent F, Araújo NH, Gully D, Klopp C, Giraud E, Arrighi JF. 2022. A mutant-based analysis of the establishment of Nod-independent symbiosis in the legume *Aeschynomene evenia*. Plant Physiol. 2022 Sep 28;190(2):1400-1417. doi: 10.1093/plphys/kiac32
- Ratu STN, Teulet A, Miwa H, Masuda S, Nguyen HP, Yasuda M, et al. Rhizobia use a pathogenic-like effector to hijack leguminous nodulation signaling. Sci. Rep. 2021; 11(1):2034. doi:10.1038/s41598-021-81598-6.

- Rightmyer AP, Long SR (2011) Pseudonodule formation by wild-type and symbiotic mutant Medicago truncatula in response to auxin transport inhibitors. Mol Plant Microbe Interact 24: 1372–1384
- Ren H, Gray WM. 2015. SAUR Proteins as Effectors of Hormonal and Environmental Signals in Plant Growth. Mol Plant. 2015 Aug;8(8):1153-64
- Ribeiro A, Akkermans ADL, van Kammen A, Bisseling T, Pawlowski K. 1995. A nodule-specific gene encoding a subtilisin-like protease is expressed in early stages of actinorhizal nodule development. Plant Cell 7: 785–794.
- Roy S, Liu W, Nandety RS, Crook A, Mysore KS, Pislariu CI et al. Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. Plant Cell. 2020; 32:15-41.
- Sadowsky MJ, Tully RE, Cregan PB, Keyser HH. Genetic diversity in *Bradyrhizobium japonicum* serogroup 123 and its relation to genotype-specific nodulation of soybean. Appl Environ Microbiol. 1987; 53:2624-30.
- Sambrook J, Fritsch EF, Maniatis TA.Molecular cloning: a laboratory manual, 2nd ed. 1989. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Schaller A, Stintzi A, Rivas S, Serrano I, Chichkova NV, Vartapetian AB, Martínez D, Guiamét JJ, Sueldo DJ, van der Hoorn RAL, Ramírez V, Vera P.2018. From structure to function - a family portrait of plant subtilases. New Phytol. 2018 May;218(3):901-915
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, et al. Genome sequence of the palaeopolyploid soybean. Nature. 2010 Jan 14;463(7278):178-83.
- Shaar-Moshe L, Brady SM. 2023. SHORT-ROOT and SCARECROW homologs regulate patterning of diverse cell types within and between species. New Phytol. 2023 Mar;237(5):1542-1549.

- Singh S, Katzer K, Lambert J, Cerri M, Parniske M. 2014. CYCLOPS, a DNA-binding transcriptional activator, orchestrates symbiotic root nodule development. Cell Host Microbe. 15(2):139-52. doi: 10.1016/j.chom.2014.01.011
- Songwattana P, Chaintreuil C, Wongdee J, Teulet A, Mbaye M, Piromyou P, Gully D, Fardoux J, Zoumman AMA, Camuel A, Tittabutr P, Teaumroong N, Giraud E. 2021. Identification of type III effectors modulating the symbiotic properties of *Bradyrhizobium vignae* strain ORS3257 with various *Vigna* species. Sci Rep. 2021 Mar 1;11(1):4874. doi: 10.1038/s41598-021-84205-w.
- Svistoonoff S, Laplaze L,AuguyF, Runions J,DuponnoisR, Haseloff J, FrancheC, Bogusz D. 2003. cg12 expression is specifically linked to infection of root hairs and cortical cells during *Casuarina glauca* and *Allocasuarina verticillata* actinorhizal nodule development. Molecular Plant-Microbe Interactions 16: 600– 607.
- Staehelin C, Krishnan HB. Nodulation outer proteins: double-edged swords of symbiotic rhizobia. Biochem. J. 2015; 470:263-274.
- Takeda N, Sato S, Asamizu E, Tabata S, Parniske M. 2009. Apoplastic plant subtilases support arbuscular mycorrhiza development in Lotus japonicus. Plant Journal 58: 766–777.
- Teulet A, Camuel A, Perret X, Giraud E. The versatile roles of Type III secretion systems in rhizobium-legume symbioses. Annu Rev Microbiol. 2022; 76:45-65.
- Teulet A, Busset N, Fardoux J, Gully D, Chaintreuil C, Cartieaux F. et al. The rhizobial type III effector ErnA confers the ability to form nodules in legumes. Proc. Natl. Acad. Sci. USA. 2019; 116:21758-21768.
- Tsuda K, Somssich, I.E. 2015. Transcriptional networks in plant immunity. New Phytolologist 206:932-947
- Vernié T, Camut S, Camps C, Rembliere C, de Carvalho-Niebel F, Mbengue M, Timmers T, Gasciolli V, Thompson R, le Signor C, Lefebvre B, Cullimore J, Hervé C. PUB1 interacts

with the receptor kinase DMI2 and negatively regulates rhizobial and arbuscular mycorrhizal symbioses through its ubiquitination activity in *Medicago truncatula*. Plant Physiol. 2016 Apr;170(4):2312-24. doi: 10.1104/pp.15.01694

- Vincent J. A manual for the practical study of root-nodule bacteria. Blackwell Scientific Publications. 1970; Oxford, U.K.

Supplementary information for

New insights into the T3SS-dependent symbiotic pathway between *Aeschynomene evenia* and *Bradyrhizobium* strains

Alicia Camuel[1,2], Djamel Gully[1,2], Marjorie Pervent[1,2], Albin Teulet[1,3], Nico Nouwen[1,2], Jean-François Arrighi[1,2*], Eric Giraud[1,2,*]

* Corresponding authors: eric.giraud@ird.fr, jean-francois.arrighi@ird.fr

This document includes only Table S1 to S3.

Supplementary Table S1. Plant accessions used in this study, origin and characteristics.

Species	LSTM code	Accession	Chromosome number (2n)	2C DNA content (pg)	Origin	Seed bank
A. indica	19	LSTM 19	2n=60	2.63 ± 0.02	Senegal	LSTM
A. evenia ssp. *evenia*	21	PI 225551	2n=20	0.85 ± 0.01	Zambia	USDA
	45	LSTM 45		0.87 ± 0.01	Senegal	LSTM
	57	IRRI 013015			Zimbabwe	IRRI
	75	CIAT 22700		0.86 ± 0.01	Senegal	CIAT
	76	CIAT 22838		0.83 ± 0.00	Malawi	CIAT
	78	CIAT 08426		0.86 ± 0.01	Brazil	CIAT
	79	CIAT 08251		0.87 ± 0.01	Brazil	CIAT
	80	CIAT 08232		0.82 ± 0.01	Brazil	CIAT
	81	CIAT 08261		0.83 ± 0.01	Brazil	CIAT
	90	ATF 3087		0.83 ± 0.01	Argentina	AusPGRIS
	101	CPI 090919		0.85 ± 0.02	Mexico	AusPGRIS
	108	CIAT 22901		0.93 ± 0.01	Colombia*	CIAT
	109	unknown				
	127	ILRI 013262			Kenya	ILRI
	143	ILRI 17651			Nigeria	ILRI
	144	ILRI 16625			Namibia	ILRI
	172	IRRI 013058			Madagascar	IRRI
	202	CIAT 22458		0.86 ± 0.00	Argentina	CIAT
	218	ILRI 15713			Chad	ILRI

* Location of the germplasm bank

Supplementary Table S2. Strains and mutants used in this study, origin and characteristics.

	Species	Strain	Supergroup	Characteristic
Bradyrhizobium	*lablabi*	CCBAU23086^{T}	*B. jicamae*	
Bradyrhizobium	sp.	LMG09283	*B. japonicum*	
Bradyrhizobium	*centrosematis*	LMG29515^{T}	*B. japonicum*	
Bradyrhizobium	*icense*	LMTR13^{T}	*B. jicamae*	
Bradyrhizobium	sp.	NAS96.2	*B. elkanii*	
		NAS96.2Ω*rhcN*		Camuel et al. 2023
		NAS96.2Ω*sup3*		Camuel et al. 2023
Bradyrhizobium	*vignae*	ORS3257	*B. japonicum*	
		ORS3257Δ*T3SS*		Teulet et al. 2019
		ORS3257Δ*ernA*		Teulet et al. 2019
		ORS3257::*sup3*		This study
		ORS3257Δ*ernA::sup3*		Camuel et al. 2023
Bradyrhizobium	sp.	ORS86	*B. elkanii*	
		ORS86Ω*rhcN*		
		ORS86Ω*ernA*		
		ORS86Ω*sup3*		
Bradyrhizobium	*diazoefficiens*	SEMIA5080	*B. japonicum*	
Bradyrhizobium	sp.	STM3561	*B. japonicum*	
Bradyrhizobium	*japonicum*	USDA123	*B. japonicum*	
Bradyrhizobium	*elkanii*	USDA61	*B. japonicum*	
		USDA61Ω*rhcJ*		Okazaki et al. 2009
		USDA61Ω*ernA*		Teulet et al. 2019
Bradyrhizobium	*elkanii*	USDA76^{T}	*B. japonicum*	
Bradyrhizobium	*archetypum*	WSM1744^{T}	*B. jicamae*	
		WSM1744Δ*T3SS*		Camuel et al. 2023
		WSM1744Δ*sup3*		Camuel et al. 2023

Supplementary Table S3. Summary of raw Illumina sequencing and filtered reads after trimming and alignment of reads to *Aeschynomenae evenia* v1 genome in each library.

Samples	Raw reads	Filter reads (%)	Unique mapped reads (%)
wt-2dpi	76 224 020	99,93	95,53
	75 414 475	99,97	95,57
	75 301 981	99,96	95,65
wt-3dpi	81 757 331	99,96	95,18
	84 001 924	99,96	95,60
	71 964 183	99,96	94,22
wt-4dpi	75 963 631	99,94	95,21
	82 700 243	99,90	93,10
	70 577 380	99,96	95,32
wt-9dpi	83 600 424	99,95	95,36
	66 620 366	99,98	95,78
	96 936 454	99,96	95,39
Δ*ernA* -2dpi	79 003 580	99,98	95,55
	73 570 719	99,95	93,98
	78 359 124	99,91	94,84
Δ*ernA* -3dpi	76 063 076	99,86	93,84
	75 880 612	99,94	95,12
	84 584 498	99,97	94,70
Δ*ernA* -4dpi	69 428 305	99,91	95,30
	81 364 114	99,91	95,07
	73 885 225	99,95	95,37
Δ*ernA* -9dpi	82 542 414	99,97	93,40
	81 676 285	99,94	95,83
	86 573 612	99,89	94,17

Chapitre 3

Caractérisation fonctionnelle des ET-Nods (projet ANR "ET-Nod")

1. Contexte de l'étude

Pour aller plus loin dans la caractérisation du processus Nod-indépendant, T3SS-dépendant, un des objectifs de ma thèse, qui s'inscrit dans le cadre du projet ANR "ET-Nod", a été d'avancer dans la compréhension du mode d'action des ET-Nods en essayant d'identifier leurs cibles chez la plante hôte après leur translocation. La caractérisation fonctionnelle d'ErnA a été initiée avant mon travail de thèse et a permis de montrer qu'il est adressé au noyau et qu'il interagit avec les acides nucléiques (Teulet et al., 2019). Par ailleurs, dans l'article 1, un certain nombre de motifs de la protéine ErnA ont pu être identifiés comme requis pour son activité. Enfin, l'identification de la plante modèle *A. evenia* N21 présentée dans l'article 2 nous permet maintenant d'utiliser des outils adaptés pour l'étude fonctionnelle des ET-Nods *in planta*. Diverses stratégies sont alors envisagées pour identifier les cibles ADN ou ARN d'ErnA ainsi que ses interactants protéiques en collaboration avec deux autres équipes faisant parties du projet ANR "ET-Nod". L'idée de ces approches multiples est d'essayer de faire émerger une liste réduite de candidats communs, qui devront par la suite être validés fonctionnellement.

Ainsi, le groupe de Laurent Deslandes met en place des approches *in vivo* de ChIP-Seq et RIP-Seq[120] et des approches *in vitro* de DAP-Seq[121] pour identifier des cibles potentielles nucléaires ainsi qu'une approche de marquage de proximité (BioID) pour identifier l'interactome d'ErnA. Le groupe de Peter Mergaert va quant à lui mener des approches de Co-IP[122] et de double hybride en levure[123] avec *a priori*, pour tester l'interaction d'ErnA avec les SUMO identifiées chez *A. evenia* N21 (Ae07g02130, Ae0108060, Ae07g04859) mais aussi avec certains déterminants canoniques de la voie de signalisation symbiotique commune (AePOLLUX, AeCCaMK, AeCYCLOPS, AeNSP2 et AeNIN). Enfin de mon côté, j'ai mené durant ma dernière année de thèse des analyses transcriptomiques comparatives par RNA-seq de racines transformées d'*A. evenia* N21 exprimant *ernA* ou non pour identifier les gènes impactés par cet ET-Nod. Pour cela, il a été décidé d'utiliser un promoteur inductible à la dexaméthа-

[120] Chromatin Immunoprecipitation sequencing (ChIP-Seq) ; RNA Immunoprecipitation sequencing (RIP-Seq). Techniques *in vivo* permettant l'identification des sites de liaison de l'ADN ou de l'ARN pour une protéine donnée.

[121] DNA affinity purification sequencing. Technique *in vitro* similaire au ChIP-Seq qui ne requiert pas l'utilisation d'anticorps.

[122] Co-immunoprecipitation. Technique de biochimie *in vitro* utilisée pour identifier une interaction protéine-protéine en utilisant des anticorps spécifiques pour "capturer" indirectement les protéines liées à une protéine ciblée.

[123] En anglais, "yeast-two-hybrid". Technique de biologie moléculaire utilisée pour identifier des interactions protéine-protéine et protéine-ADN en système hétérologue : la levure.

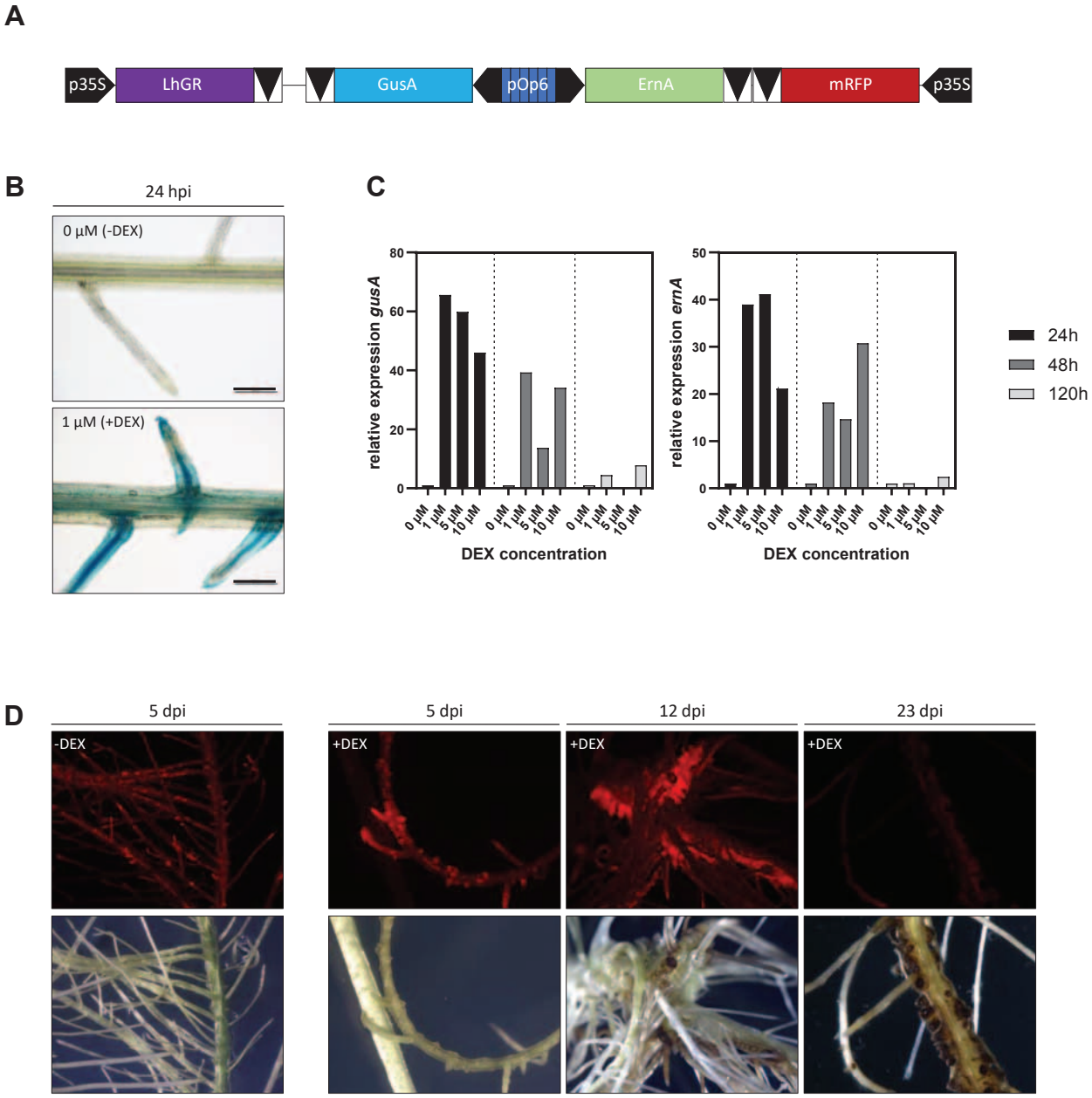


Figure 63. Le système pDEX permet de contrôler le niveau d'expression d'*ernA* par le dexaméthasone dans des racines transgéniques d'*A. evenia*.

(**A**) Construction génétique utilisée pour la transformation transitoire de racines d'*A. evenia* avec le promoteur pDEX-*ernA*. Le facteur de transcription codé par le gène *LhGR* ne peut rentrer dans le noyau et se lier au promoteur bidirectionnel pOp6 qu'en présence de dexaméthasone. Une fois le promoteur activé, il induit simultanément l'expression des gènes *ernA* et *gusA*. (**B**) Racines transformées d'*A. evenia* observées en lumière blanche, 24h après induction au dexaméthasone et incubées dans une solution de coloration X-gluc ; barre d'échelle : 500 µm. (**C**) Expression relative des gènes *gusA* (gauche) et *ernA* (droite) dans des racines transformées avec le plasmide pDEX-*ernA* avec ou sans induction au dexaméthasone. Les valeurs ont été normalisées par l'expression à 0 µM. (**D**) Phénotype de racines transformées d'*A. evenia* avec le plasmide pDEX-*ernA* avec ou sans induction au dexaméthasone. Une extinction de la mRFP et des tissus nécrotiques sont observables à des temps plus tardifs (23 dpi), dpi : day post-induction.

sone pour contrôler le timing d'expression d'*ernA* et ainsi détecter à des temps précoces ses cibles potentielles.

2. Avancements dans la caractérisation fonctionnelle d'ErnA

2.1. Recherche des cibles précoces régulées transcriptionnellement par ErnA

Dans le chapitre 2, il a été réalisé une analyse transcriptomique chez des plantes ayant été inoculées par la souche sauvage ORS3257 ou par le mutant de délétion ORS3257Δ*ernA*, dans le but d'identifier des gènes candidats sous le contrôle d'ErnA. Cependant cette stratégie utilise le système naturel du T3SS de la bactérie présente dans le milieu pour injecter un cocktail d'effecteurs (avec ou sans ErnA) qui peuvent aussi influer l'expression génique de la plante. De plus, on suppose que la réponse de la plante à l'injection d'ErnA doit par ailleurs être extrêmement localisée et donc sûrement très diluée lorsqu'on fait une analyse transcriptomique sur une racine entière. Il a donc été décidé de mener une nouvelle analyse transcriptomique qui cette fois se ferait à partir de racines hairy-root exprimant seulement *ernA* afin d'éviter des éventuelles interférences avec d'autres T3Es ou d'autres composants de la bactérie. Par ailleurs, *ernA* a été mis sous le contrôle d'un promoteur inductible de façon à contrôler l'activation de son expression.

2.1.1. Développement d'un système d'expression inductible d'*ernA* chez *A. evenia*

Pour identifier les cibles d'ErnA à des temps précoces, il a été décidé d'utiliser un promoteur inductible à la dexaméthasone[124] pour contrôler son expression : le système pOp6-LhGR (Samalova et al., 2005) que nous appellerons ici pDEX. Il est constitué d'un promoteur bidirectionnel qui active à la fois le gène rapporteur *gusA* dans un sens, ainsi que notre gène d'intérêt dans l'autre sens (figure 63A). Ce système a été utilisé avec succès chez une variété de plantes modèles (Schena et al., 1991; Borghi, 2010; Samalova et al., 2019), mais n'a jamais été développé pour une étude chez *Aeschynomene*. Selon les plantes étudiées, les concentrations en dexaméthasone peuvent être variables pour induire l'expression du transgène et il est important de vérifier que le phénotype racinaire attendu est induit. Il a donc fallu mettre au point un protocole pour déterminer les conditions expérimentales favorables au système pDEX chez *A. evenia* N21. La méthodologie utilisée a été de cloner *ernA* dans le vecteur inductible pDEX-GW-SBP qui nous a été fourni par le groupe de Laurent Deslandes. Dans cette construction, il a été rajouté le gène *mRFP* sous le contrôle d'un promoteur 35S pour pouvoir repérer les racines transformées et faciliter la sélection. Après obtention de notre construction dans

[124] Glucocorticoïde.

le plasmide pDEX[125], une expérience préliminaire a pu être menée avec un stagiaire de Master 1 Vetea Jacot que j'ai encadré. Des tests d'induction ont été réalisés sur des plantes d'*A. evenia* N21 âgées de 18 jours en utilisant différentes concentrations de dexaméthasone et en suivant l'expression d'*ernA* sur 5 jours. L'activation du promoteur bidirectionnel par ce glucocorticoïde a pu être vérifiée dans un premier temps en utilisant le gène rapporteur *gusA* par coloration au X-gluc des racines transformées (figure 63B). Des RT-qPCR réalisées sur les racines transformées ont confirmé qu'il est possible d'exprimer fortement *ernA* avec un pic d'expression autour de 1 µM de dexaméthasone 24h après induction (figure 63C). Des observations réalisées sur des temps plus tardifs (12 jours après induction) montrent également que les racines transformées avec ce système pDEX-*ernA* présentent le phénotype attendu ; c'est-à-dire l'induction de structures de type nodulaires ainsi que des racines secondaires dont le développement s'est arrêté (figure 63D).

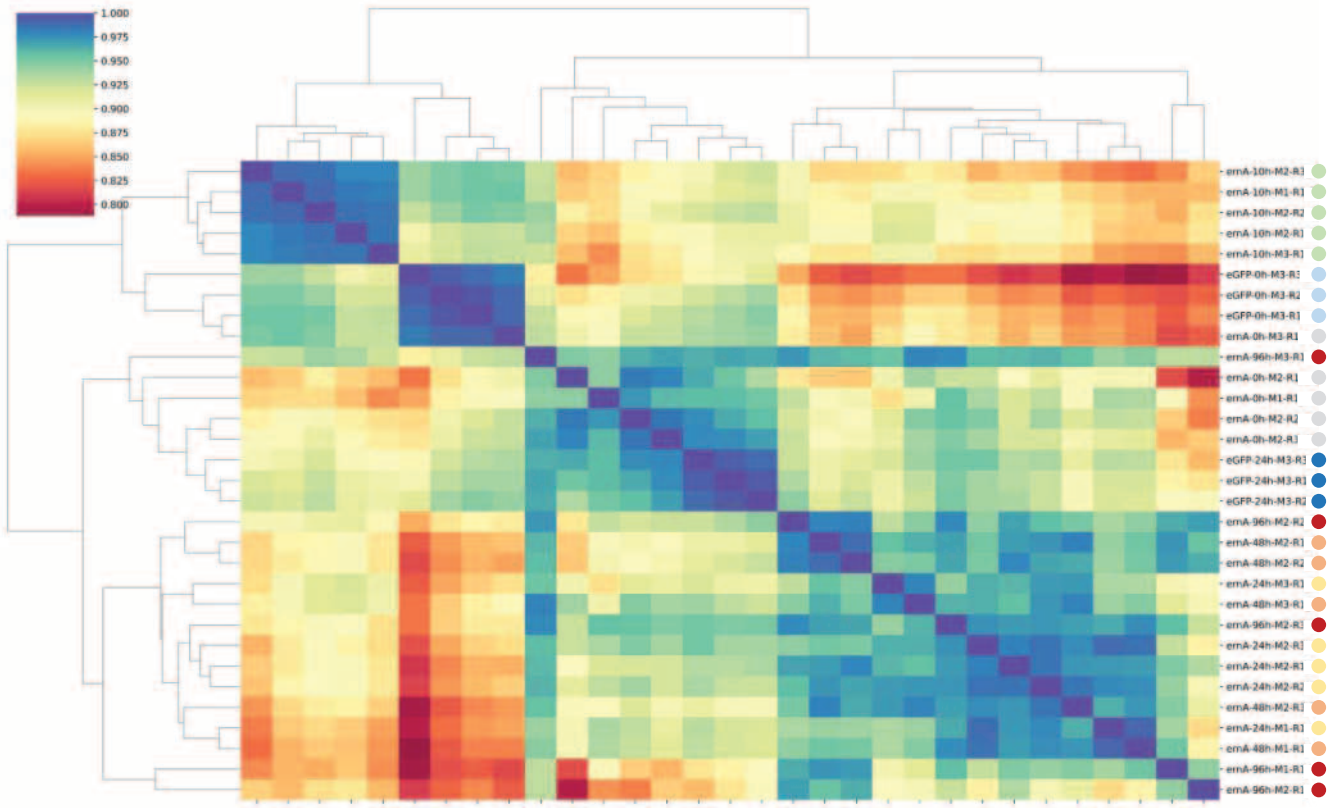
Il a pu être suivi par RT-qPCR sur les conditions 0-1µM/24h l'expression de différents gènes connus de la voie de signalisation symbiotique et des gènes candidats identifiés dans l'analyse RNA-seq faite précédemment (cf. article 2). Aucun des gènes testés de la voie de signalisation symbiotique semble être induit 24h après induction d'*ernA* ; cependant la subtilase (Ae03g08090), qui avait été trouvée la plus fortement exprimée à 48h dans la condition sauvage par rapport au mutant Δ*ernA*, est retrouvée fortement induite après ajout de dexaméthasone sur les racines transformées. Ce résultat prometteur nous conforte dans l'idée que cette approche est pertinente pour identifier les cibles sous le contrôle transcriptionnel d'ErnA. L'obtention du matériel final envoyé au séquençage a été réalisée avec le soutien technique d'une étudiante en césure Melissa Simonnet.

2.1.2. Analyse différentielle des gènes potentiellement sous le contrôle d'ErnA

La comparaison du transcriptome de racines exprimant ou pas *ernA* grâce à ce système pDEX a ensuite été réalisée après avoir déterminé la concentration de dexaméthasone et les temps après induction les plus appropriés pour notre étude. Notre choix s'est porté sur une concentration de 1 µM avec des prélèvements à 10h, 24h, 48h et 96h après induction sur des racines transformées d'*A. evenia* N21 âgées de 18 jours. Nous avons ajouté à cette analyse le témoin pDEX-*NLS-eGFP* à 0h et 24h après induction pour déterminer les gènes qui pourraient être exprimé de manière non-spécifique à ErnA *via* le système inductible à la dexaméthasone. Le choix du contrôle s'était initialement porté sur une forme mutée du pDEX-*ernA** rendant la protéine non fonctionnelle par induction d'un codon STOP prématuré. Cependant, en raison de difficultés rencontrées lors de la mutagénèse dirigée, nous

[125] Plasmide final : *pDEX-ernA-35S-mRFP*.

A



B

Condition	Up-genes	Down-genes	Ratio up/down
ErnA 10h VS 0h	453	108	4,19
ErnA 24h VS 0h	568	28	20,29
ErnA 48h VS 0h	470	22	21,36
ErnA 96h VS 0h	528	20	26,40
eGFP 24h VS 0h	152	61	2,49

C

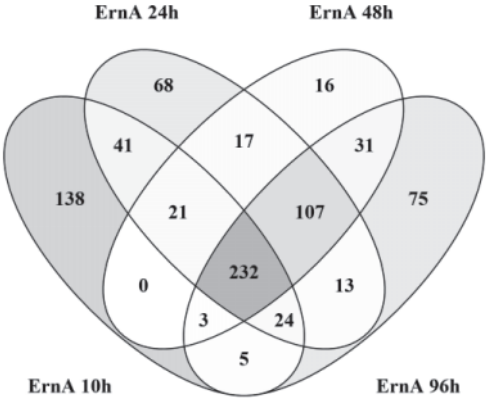


Figure 64. Comparaison des transcriptomes de racines transformées exprimant ou pas *ernA*.
(**A**) Heat-map représentant le niveau de transcrits pour chacune des conditions analysées. La construction pDEX-*ernA* ou pDEX-*NLS-eGFP* (contrôle) ainsi que le numéro de l'expérimentation (répétition biologique) où sont issus les échantillons (M1, M2 ou M3) sont associés à une couleur de pastille dans la colonne de droite, R indique le numéro de la répétition technique (1, 2 ou 3). (**B**) Nombre de gènes surexprimés ou réprimés après induction au dexaméthasone du pDEX-*ernA* ou pDEX-*NLS-eGFP* par rapport à la condition non induite. (**C**) Diagramme de Venn illustrant le nombre de gènes différentiellement exprimés entre les conditions pDEX-*ernA* avant et après induction au dexaméthasone, et après retrait des gènes différentiellement exprimés dans la condition témoin pDEX-*NLS-eGFP*. Les gènes différentiellement exprimés (DEG) ont été définis en utilisant une FDR ≤ 0,05 et un log2 (FC) ≥ 2. Le diagramme de Venn a été obtenu sur le logiciel en ligne https://bioinfogp.cnb.csic.es/tools/venny.

avons décidé d'utiliser un témoin exprimant une protéine eGFP adressée au noyau[126] comme l'est ErnA (construction réalisée par le groupe de Laurent Deslandes).

Le niveau de reproductibilité de chaque répétition a été analysé et l'ensemble des conditions sélectionnées présente des niveaux de transcrits comparables à l'exception des échantillons de la condition 96h qui apparaissent moins associés les uns aux autres (figure 64A). Cela n'est pas forcément surprenant puisque à des temps plus tardifs d'expression d'*ernA*, un grand nombre de perturbations développementale peut être observé et la transformation transitoire par *A. rhizogenes* ne permet pas de contrôler le nombre d'insertion et donc le niveau d'expression du transgène qui peut varier d'une plante à l'autre. Les gènes différentiellement exprimés entre des racines transformées non induites et des racines transformées induites à la dexaméthasone à chaque temps de prélèvement ont été identifiés en utilisant une FDR[127] de 0,05 et un FC[128] de log2 ≥ 2.

Il apparaît de manière globale que l'induction de l'expression d'*ernA* provoque une importante surexpression, à des niveaux élevés, plutôt qu'une répression de gènes (figure 64B). C'est donc une réponse beaucoup plus forte qui est observée en réponse à l'expression d'*ernA* que celle qui a pu être obtenue lors de l'analyse conduite précédemment sur la bactérie ORS3257 et son dérivé Δ*ernA*. Cela suggère qu'ErnA pourrait agir sur le niveau d'expression de certains facteurs de transcriptions (FTs) généraux qui vont à leur tour activer l'expression de multiples gènes en aval. Cette hypothèse peut être appuyée par le grand nombre de gènes surexprimés et retrouvés communs aux quatre conditions analysées et qui s'élève à 232 (figure 64C). Dans ce sens, on retrouve aussi parmi les gènes les plus fortement exprimés un certain nombre de facteurs de transcriptions dont nous discuterons ci-dessous et qui nous semblent être les candidats les plus intéressants. Nous nous sommes basés pour notre analyse sur les gènes différentiellement exprimés à 24h après induction car il est apparu qu'à la condition 10h, les gènes spécifiquement exprimés (138) présentent des log2 de FC à la limite du seuil imposé et ne correspondent pas à des candidats notables. Enfin, pour la suite de l'analyse, nous avons supprimé les gènes communs aux conditions *pDEX-ernA* et *pDEX-eGFP*.

2.1.3. De nombreux facteurs de transcription sont surexprimés en réponse à *ernA*

Comme précédemment observé lors de l'analyse transcriptomique de la bactérie ORS3257 et de son mutant Δ*ernA*, très peu de gènes symbiotiques déjà décrits sont retrouvés dans cette analyse *pDEX-ernA* (figure 65). Néanmoins, l'un des gènes identifié parmi les plus fortement exprimés dans la condition *pDEX-ernA* code pour le facteur de transcription NF-YC (Ae06g34470), qui se trouve être une

[126] Plasmide final : *pDEX-3HA-NLS-eGFP-35S-mRFP*. La fusion 3HA a été ajoutée pour utiliser ce contrôle dans d'autres approches telle que le Turbo-ID.
[127] False Discovery Rate.
[128] Fold Change.

Symbiotic-related genes
NF-YC

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae06g34470	6,82	9,57	10,43	10,67

Nodulins

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae06g09080	2,76	4,95	4,66	4,42
Ae04g04380	-	3,37	3,35	2,79

Subtilisin-like protease

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae03g08090	4,07	5,85	5,74	5,44

Signalling pathways Log2(FC)>4
RLK, RLP, protein kinase

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae07g01650	9,36	11,28	9,58	9,10
Ae09g24180	9,38	10,59	9,30	9,46
Ae03g22520	7,78	10,41	10,64	11,60
Ae10g23550	9,00	10,30	7,63	7,89
Ae05g00990	9,53	8,48	6,65	6,40
Ae07g06130	7,53	8,40	5,78	6,37
Ae01g06420	5,87	8,14	7,57	8,24
Ae08g26180	9,13	8,19	5,53	5,28
Ae05g28230	6,31	7,97	6,80	6,49
Ae06g34140	7,10	7,84	6,26	5,63
Ae09g24280	6,37	7,64	6,44	6,91
Ae08g03190	6,42	7,48	6,12	5,79
Ae02g20080	5,33	7,06	5,21	5,29
Ae08g25920	7,09	7,04	4,39	5,55
Ae05g27150	4,71	6,87	5,25	5,07
Ae04g02260	6,15	6,73	6,23	5,67
Ae08g07470	4,64	6,61	5,85	5,66
Ae05g08460	4,29	6,17	5,34	5,76
Ae02g20020	3,33	6,03	3,66	4,28
Ae04g10870	5,76	5,90	3,26	3,06
Ae03g22490	3,60	5,82	5,81	6,11
Ae06g36470	5,43	5,77	4,41	4,41
Ae06g36600	3,33	5,74	3,93	3,88
Ae09g24160	3,69	5,67	5,10	5,70
Ae08g08180	4,77	5,66	4,33	4,60
Ae04g09960	4,40	5,50	5,32	5,48
Ae09g24250	4,26	5,49	4,36	4,62
Ae04g02860	4,67	5,37	4,08	4,72
Ae06g09820	3,49	5,26	6,24	6,75
Ae08g08190	3,40	5,21	4,00	4,46
Ae09g24260	4,74	4,88	3,80	3,96
Ae09g24140	2,45	4,29	3,23	3,49
Ae03g02350	3,79	4,17	-	2,45
Ae09g24220	3,66	4,08	2,86	3,00
Ae05g36820	2,98	4,03	2,85	2,48
Ae02g20010	-	4,01	3,69	3,73

Transcriptionnal factors
bHLH family

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae01g26780	6,27	6,40	6,19	6,03
Ae04g13790	3,88	4,27	3,06	3,21
Ae06g10870	-	2,58	-	-

WRKY family

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae07g13440	5,16	6,22	5,06	5,09
Ae08g06820	6,04	6,21	4,75	4,73
Ae02g19030	5,29	6,14	5,17	5,70
Ae08g08320	4,66	4,98	4,33	4,16
Ae08g04850	-	3,02	2,62	2,05
Ae03g30330	2,03	3,00	3,28	3,79
Ae09g10580	-	2,09	2,32	2,46

ERF family

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae04g04370	3,25	4,01	4,95	4,85
Ae01g03740	-	2,27	-	-
Ae09g17140	-	2,14	-	-

MADS family

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae09g24320	3,52	5,10	4,00	3,45
Ae09g24330	3,18	4,32	3,51	3,29

Others

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae07g22680	3,94	4,31	-	-
Ae04g26600	-	4,05	-	3,46
Ae10g21570	2,20	3,08	2,85	3,40
Ae07g09790	-	2,81	2,18	2,39
Ae05g10290	2,68	2,49	2,61	3,14
Ae02g01090	-	2,34	-	-
Ae04g35720	2,19	2,34	-	-
Ae06g01880	-	2,27	-	-
Ae10g15850	2,35	2,10	-	-

Flavonoids-related genes
Enzymes from flavonoids synthesis.

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae04g36280	7,35	9,09	9,46	9,31
Ae06g16470	5,24	7,62	8,69	9,25
Ae05g25050	-	4,27	-	3,52
Ae05g11080	3,39	4,26	4,19	3,99
Ae05g11540	2,91	3,75	3,68	3,58
Ae05g04910	2,80	3,42	3,04	2,63
Ae10g01370	-	3,10	2,94	3,02
Ae10g20280	-	2,80	2,69	2,14
Ae05g11090	-	2,51	2,20	2,60
Ae05g11160	-	2,49	2,27	-
Ae08g09700	2,78	2,48	-	2,44
Ae03g16280	-	2,47	-	-
Ae02g20260	-	2,38	-	-
Ae05g11110	-	2,24	2,23	2,04

Defense-related genes
Chitinase

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae10g17060	7,29	7,82	7,64	7,82
Ae05g03570	-	3,68	3,76	3,48
Ae01g01290	2,97	2,53	3,61	3,77
Ae04g32820	2,35	2,20	-	-

Peroxidase

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae09g21100	2,52	3,86	2,51	2,14
Ae01g13550	-	2,82	2,75	2,05
Ae01g23340	2,49	2,35	-	-

TIR-NBS-LRR

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae02g22200	-	3,68	6,18	6,63
Ae01g16290	-	2,85	2,26	2,50

Others

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae06g12000	4,08	6,79	7,52	7,87
Ae01g29070	-	3,57	3,52	3,30
Ae01g29080	-	3,45	3,53	3,45
Ae03g04800	2,58	2,86	2,60	2,80
Ae05g03410	-	2,34	-	-

Cell wall-related genes
Pectinesterase, endoglucanase etc.

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae02g19860	-	8,21	8,96	8,13
Ae02g19870	-	8,15	9,58	8,21
Ae04g16490	-	7,19	7,26	7,46
Ae01g29230	5,04	7,08	7,58	7,66
Ae04g12670	4,71	6,22	5,74	6,27
Ae04g35860	2,50	4,29	4,56	4,74
Ae06g11170	3,67	4,28	-	2,18
Ae09g20760	4,53	4,08	4,06	3,89
Ae08g05050	4,46	3,70	2,82	3,37
Ae04g35850	-	3,42	3,98	4,30
Ae05g00910	2,07	3,24	5,08	5,54
Ae05g01410	-	2,51	2,50	2,07
Ae05g00920	-	2,35	2,18	2,08
Ae05g26800	2,12	2,07	-	-

Actin-binding protein

Gene ID	Log2 (FC)			
	10 hpi	24 hpi	48 hpi	96 hpi
Ae04g26790	5,88	4,62	5,12	4,84
Ae03g04830	-	2,52	2,34	2,15

Figure 65. Analyse RNA-seq entre des racines d'*A. evenia* N21 transformées avec la construction pDEX-*ernA*, avec ou sans induction à la dexaméthasone.
Identifiant et catégorie des principaux gènes différentiellement exprimés (DEG) entre la condition induite et non induite à 10, 24, 48 et 96 heures après induction (hpi) ; et après retrait des gènes identifiés surexprimés en présence de dexaméthasone dans la condition contrôle *pDEX-eGFP*. Un gradient de couleur du vert clair au vert foncé indique le Fold Change (FC) des gènes surexprimés dans les racines de la condition *pDEX-ernA* induite par rapport aux racines *pDEX-ernA* non-induites.

sous-unité du complexe NF-Y[129] impliqué dans la signalisation symbiotique chez les légumineuses (Zanetti et al., 2017). Le complexe NF-Y est composé de trois sous-unités différentes : NF-YA, NF-YB et NF-YC (Baudin et al., 2015). NF-YA, initialement identifié chez *M. truncatula* (Combier et al., 2006), est capable de se lier à l'ADN et s'associe aux deux autres sous-unités NF-YB et NF-YC. Ces trois composants du complexe NF-Y reconnaissent spécifiquement des boîtes CCAAT retrouvées dans les promoteurs de leurs gènes cibles et interviennent à des temps précoces de la nodulation dont notamment l'infection par les rhizobiums et l'organogenèse du nodule (Zanetti et al., 2017). De manière intéressante, il a été montré que NF-YC est localisé au noyau et qu'il participe chez *Phaseolus vulgaris* au développement nodulaire en activant la division des cellules du cortex (Zanetti et al., 2010). De plus, NF-Y est connu pour être une cible directe de NIN et pour activer un certain nombre de gènes requis pour l'organogenèse nodulaire (Soyano et al., 2013). Ce complexe protéique joue donc un rôle majeur dans la régulation développementale conduisant à la transition de l'identité cellulaire d'une racine à un nodule. La surexpression observée de *AeNF-YC* en réponse à l'expression d'*ernA* est donc un résultat fort de cette analyse transcriptomique. La modification de NF-YC par SUMOylation promeut par ailleurs la formation du complexe avec NF-YB chez *A. thaliana*, soulignant l'importance d'une régulation post-traductionnelle de ces sous-unités (Huang et al., 2023). Il a pu également être montré chez cette même plante, un rôle épigénétique du complexe NF-Y dans l'accessibilité de la chromatine afin de faciliter l'accès à d'autres FTs (Hou et al., 2014). Il n'est donc pas à exclure que chez les légumineuses de tels rôles puissent être conservés.

Dans un second temps, il apparaît que la famille des FTs de type WRKY est largement représentée et fortement induite après ajout de dexaméthasone (7 gènes – figure 65). De manière intéressante, le gène le plus surexprimé à 24h après induction (Ae07g13440) est également celui qui est apparu parmi les plus surexprimés 2 jours après inoculation dans l'analyse transcriptomique de l'article 2. Cette coïncidence qui n'en est peut-être pas une, place ce gène en haut de la liste des cibles potentielles d'ErnA. Certains WRKYs ont été montrés intervenir dans divers processus biologiques tels que la résistance aux stress environnementaux ou le développement des plantes et apparaissent donc être des régulateurs centraux de nombreuses voies de régulation (Wang et al., 2023). Récemment, il a été montré que MtWRP1[130] régule positivement la nodulation chez *M. truncatula* soulignant un potentiel rôle de cette famille de FTs dans la symbiose rhizobium-légumineuse (Chen et al., 2022).

L'analyse transcriptomique des gènes différentiellement exprimés en réponse à l'induction d'*ernA* a également révélé une nouvelle catégorie de FTs appartenant à la famille des bHLH[131] et n'apparaissant pas dans l'étude de l'article 2 (figure 65). Certains de ces FTs bHLH ont été montré

[129] Nuclear Factor- Y, famille de facteurs de transcriptions très conservée chez les eucaryotes.
[130] WRKY-Related Protein 1.
[131] En anglais, basic helix-loop-helix "hélice-boucle-hélice".

comme jouant un rôle symbiotique chez *M. truncatula*, notamment en tant que régulateur négatif de la sénescence du nodule (Deng et al., 2019) ou encore en participant à la vascularisation du primordium nodulaire (Godiard et al., 2011). A noter également qu'une analyse transcriptomique de type "single cells" pour déchiffrer de manière plus minutieuse la signalisation symbiotique a récemment été conduite par le groupe de Stig Andersen (Université d'Aarhus, Danemark). Cette étude a permis de mettre en évidence la forte expression de deux gènes *bHLH* chez certaines cellules infectées du lotier durant la symbiose. Un double mutant de ces FTs perd sa capacité de nodulation ce qui souligne l'importance de cette famille de FTs dans l'organogenèse nodulaire (Stig Andersen, communication orale durant la 15ème édition de l'ENFC[132] à Naples, 2023). Il ne serait donc pas surprenant que des FTs-bHLH puissent participer à l'organogenèse du nodule médiée par ErnA.

Par ailleurs, une analyse phylo-transcriptomique comparative chez les plantes établissant une symbiose fixatrice d'azote a permis d'identifier un certain nombre de gènes communs qui sont surexprimés durant les étapes précoces de la symbiose (Vernié et al., en préparation). Cette liste a ainsi fait ressortir des gènes déjà connus tel que NIN, mais aussi un FT de la famille des WRKY et deux FTs de type bHLH (Pierre-Marc Delaux, communication orale durant la 15ème édition de l'ENFC à Naples, 2023). Les gènes candidats de notre analyse et codant pour ces types de FTs sont donc des plus prometteurs pour être les cibles d'ErnA. A noter enfin la présence de certains FTs de type MADS-box[133] surexprimés après ajout de dexaméthasone (figure 65). Cette large famille de FTs a pu être montré comme impliqué dans des voies du développement racinaire (Shah et al., 2022) mais aussi lors de la nodulation chez le soja (Liu et al., 2015) et pourrait donc représenter des candidats à considérer.

2.1.4. Autres gènes notables

Tout comme précédemment (cf. chapitre 2 – résultats), la subtilase Ae03g08090 est retrouvée surexprimée dans cette analyse RNA-seq. Cependant, alors qu'elle est fortement exprimée à 2 dpi puis réprimée à 3 et 4 dpi dans la première analyse, cette subtilase est retrouvée surexprimée tout le long de la cinétique d'induction au dexaméthasone (figure 65). Cette forte induction d'un même gène candidat vient supporter l'idée que cette protéase pourrait jouer un rôle symbiotique chez *Aeschynomene*.

De manière intéressante, une catégorie de gènes codant pour des enzymes de la voie de biosynthèse de flavonoïdes (13) apparaissent fortement surexprimées avec deux gènes majeurs codant tous deux pour une 4-coumarate:CoA ligase (Ae04g36280 et Ae06g16470) ; le 4-Coumaroyl-CoA étant un intermédiaire moléculaire du début de la voie de biosynthèse. L'action d'ErnA sur de tels gènes pour-

[132] European Nitrogen Fixation Conference.

[133] Famille de facteurs de transcriptions initialement identifiés pour leur rôle dans le développement floral chez *A. thaliana*.

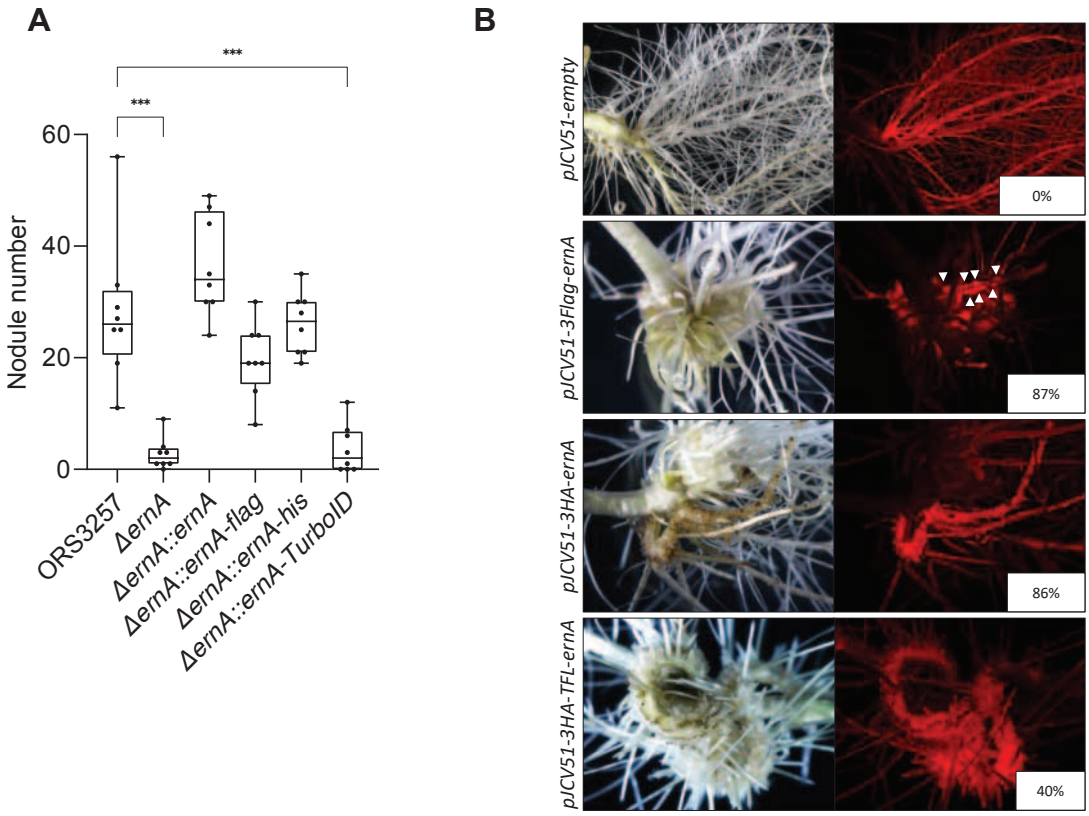


Figure 66. La fusion d'un domaine en C-ter d'ErnA peut affecter sa fonction, à l'inverse d'une fusion en N-ter.
(**A**) Nombre de nodules sur des racines d'*A. indica* 21 jours après inoculation avec la souche ORS3257 et ses différents mutants Δ*ernA* complémentés avec différentes formes taguées ou pas d'ErnA en C-ter. Box plots montrant les résultats d'une expérience réalisée sur 8 plantes. ***p < 0,001, différence significative entre la souche sauvage et chacun des mutants après un test non-paramétrique de Kruskal-Wallis. (**B**) Expression ectopique de différentes formes taguées en N-ter d'ErnA dans des racines transformées d'*A. evenia* (6 semaines après transformation), la mRFP visible en rouge par fluoresence est utilisée comme gène rapporteur de la bonne transformation des plantes. Le taux d'apparition du phénotype attendu (induction de structures de type nodulaire ou racinaire dont le développement a été avorté) est indiqué dans le rectangle blanc en bas à droite pour chacune des constructions utilisées. Le plasmide vide est utilisé comme témoin et ne présente pas le phénotype associé à l'expression d'*ernA*.

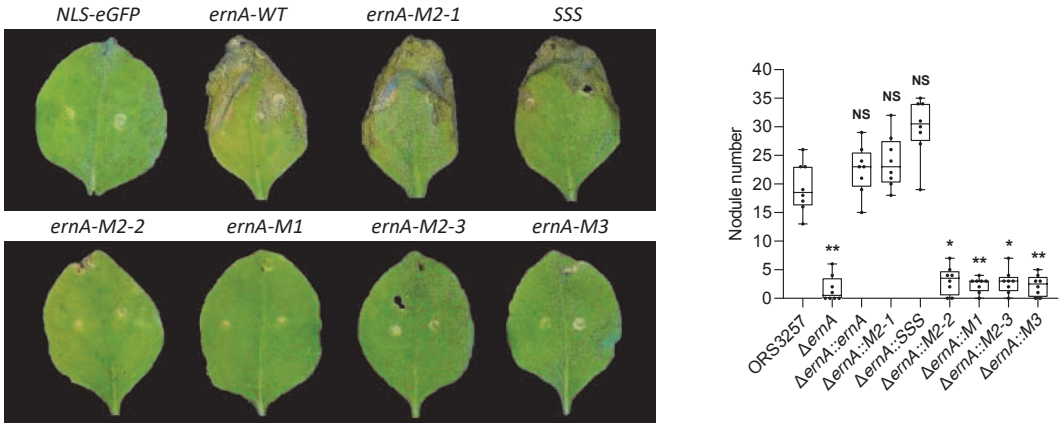


Figure 67. Induction d'une réaction hypersensible chez le tabac après agro-infiltration d'*ernA* ou de ses variants associés.
Observation des symptômes 7 jours après traitement à la dexaméthasone qui permet l'expression d'*ernA* et de ses divers variants obtenus dans le pDEX (à gauche). Une décoloration de la feuille associée à une nécrose sont visibles sur les plantes induisant une HR. Seules les formes mutées mais toujours capables de complémenter le mutant Δ*ernA* (à droite – Camuel et al., 2023) induisent une HR (à gauche). A noter qu'une légère décoloration de la feuille est visible pour la forme ernA-M2-2.

rait participer à l'activation d'une voie de signalisation requise pour la nodulation, les flavonoïdes étant des molécules impliquées dans de nombreux processus cellulaires (Mathesius, 2018). Notamment, les flavonoïdes jouent un rôle majeur sur la régulation du transport d'auxine contrôlant ainsi de nombreux aspects développementaux de la plante qui ont été décrits chez la plante modèle *A. thaliana* (Brown et al., 2001). On peut supposer que ce rôle de régulateur du transport d'auxine est conservé chez les légumineuses, l'auxine jouant un rôle central dans le développement du nodule. Dans ce sens, il a pu être montré que l'accumulation de flavonoïdes spécifiques peut conduire à une accumulation d'auxine localisée dans les cellules corticales associée à l'organogenèse nodulaire et complémenter un mutant *cre1* altéré dans la nodulation chez *M. truncatula* (Ng et al., 2015).

De nombreux gènes codant pour des enzymes liées au remodelage pariétal apparaissent fortement induits à 24h et 48h après induction. Ce sont 14 gènes qui sont surexprimés (figure 65), avec 5 particulièrement induits et codant pour des putatives inhibiteurs de pectine-esterases (Ae02g19860 et Ae02g19870), pour une pectate lyase (Ae04g16490), une polygalacturonase (Ae01g29230) et une cellulose synthase (Ae04g12670). Une telle surexpression de gènes codant diverses enzymes susceptibles de modifier la paroi cellulaire végétale, suggère qu'ErnA provoque une perturbation considérable dans le développement de la plante. En effet, la mise en place de nouveaux organes nécessite un remodelage des parois important dû aux divisions cellulaires requises pour l'organogenèse. En parallèle, une protéine capable de lier l'actine (Ae04g26790) est fortement surexprimée à chaque point de la cinétique, soulignant une éventuelle participation du cytosquelette à ces modifications développementales.

Enfin, un nombre considérable de gènes impliqués dans des voies de signalisation cellulaire peuvent être identifiés en réponse à l'ajout de dexaméthasone, dont une liste réduite est disponible dans la figure 65. Ces gènes apparaissent, pour la grande majorité, fortement exprimés à tous les points de la cinétique et codent le plus souvent pour des protéines et des récepteurs-like kinases. Cette longue liste (102 gènes au total) nous invite à penser qu'ErnA est impliqué dans l'activation de gènes centraux au cœur de nombreuses voies de signalisation. L'expression de ces gènes pourrait jouer un rôle important dans la signalisation d'autres processus que celui de l'organogenèse nodulaire, qui dans un contexte symbiotique, permettraient la reconnaissance d'autres déterminants bactériens nécessaires à la mise en place d'un nodule infecté et fixateur. Les récepteurs tels qu'Ae07g01650, Ae09g24180, Ae10g23550 et Ae10g23550 avec un log2 FC > 10 à 24h après induction, pourraient donc être des gènes clés dans la symbiose Nod-indépendante, T3SS-dépendante (figure 65).

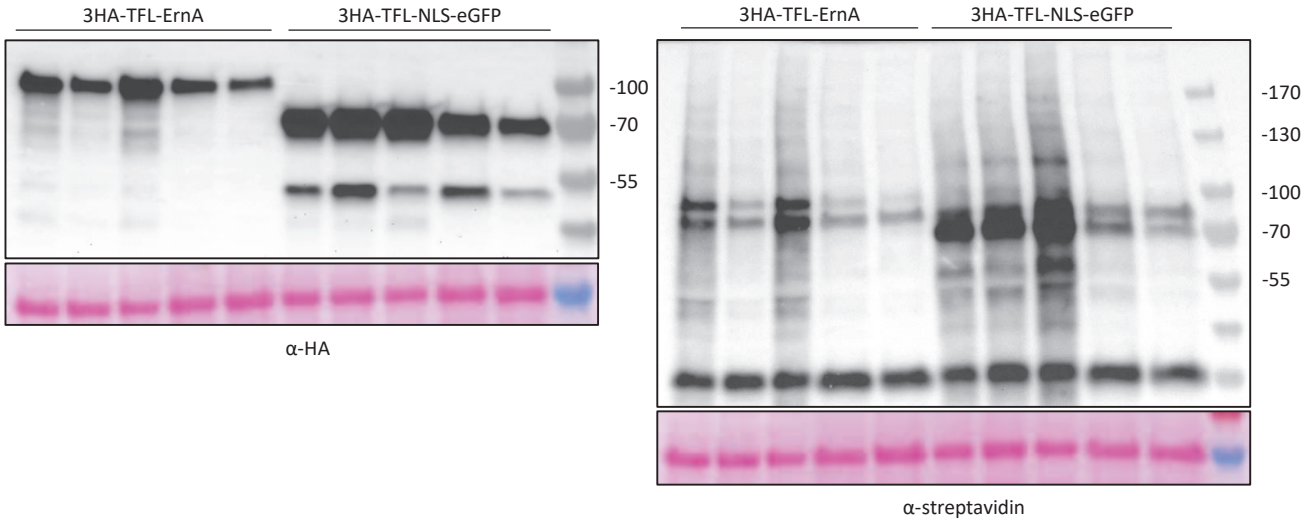


Figure 68. Validation des constructions pour le marquage de proximité des interactants d'ErnA par biotinylation.
Western-blot des protéines totales extraites de feuilles agro-infiltrées de tabac en utilisant un anticorps anti-HA (à gauche), permettant de valider l'expression des protéines taguées en N-ter, ou en utilisant un anticorps anti-streptavidine (protéine liant la biotine – à droite) permettant ainsi de valider leur capacité à biotinyler des potentielles cibles. TFL : Turbo-Full-Length.

2.2. Etablir l'interactome d'ErnA

2.2.1. Problématique de l'utilisation de formes taguées[134] d'ErnA

Les différentes approches biochimiques envisagées pour la caractérisation fonctionnelle d'ErnA nécessitent de fusionner différents tags pour la purification de la protéine, en particulier durant les tests d'immunoprécipitation (tag-6His et Flag) et de marquage de proximité (Turbo-ID). Il a donc fallu s'assurer au préalable que l'ajout d'un tag en C-terminal ou en N-terminal n'affecte pas la fonction d'ErnA. Pour cela, nous avons envisagé dans un premier temps la complémentation fonctionnelle du mutant de délétion *ORS3257ΔernA* avec des formes d'ErnA taguées en C-terminal.

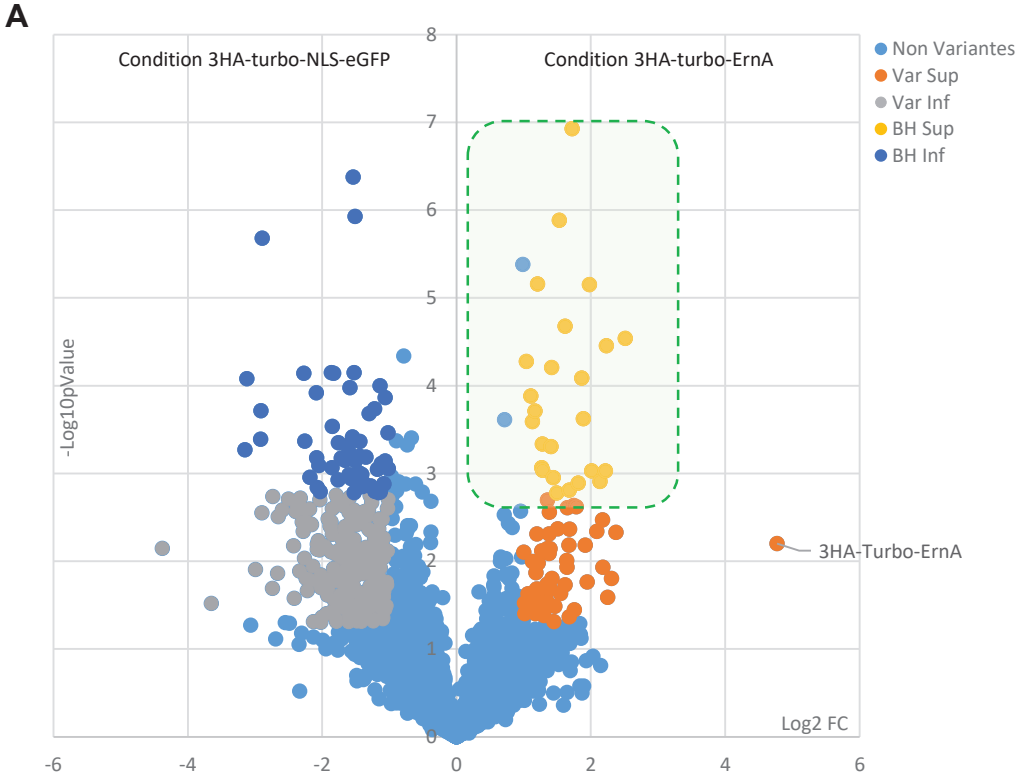
Les résultats obtenus montrent que l'ajout en C-terminal d'un tag 6His et Flag ne perturbe pas la fonction d'ErnA car on restaure une nodulation similaire à celle de la souche sauvage ORS3257. En revanche, l'ajout du tag Turbo-ID d'une taille de plusieurs kb et qui est utilisé pour l'approche de marquage de proximité, ne complémente pas le mutant de délétion (figure 66A). Comme il n'était pas possible par cette approche de tester l'effet de l'ajout d'un tag en N-terminal qui aurait empêché la sécrétion de l'effecteur, nous avons aussi testé la capacité de formes taguées en N-terminal d'ErnA à induire le phénotype attendu après expression ectopique dans *A. evenia* N21. Il a pu être ainsi vérifié que la fonctionnalité d'ErnA après ajout en N-terminal d'un tag 3Flag et 3HA (qui semblent être en fait plus appropriés pour les approches envisagées) n'est pas perturbée. La stratégie de travailler avec un tag en N-terminal semble donc être une alternative envisageable pour limiter l'effet négatif d'un tag en C-terminal mais cela implique de travailler avec des lignées transgéniques surexprimant *ernA*. Les résultats obtenus avec la construction contenant le Turbo-ID en N-terminal d'ErnA viennent confirmer ces observations (figure 66B). Il a ainsi pu être validé cette construction qui est nécessaire pour identifier l'interactome d'ErnA par une approche de marquage de proximité.

2.2.2. Recherche d'interactants potentiels par Turbo-ID sur *Nicotiana benthamiana*

L'approche de marquage de proximité par biotinylation a été envisagée pour identifier l'interactome d'ErnA chez *A. evenia* N21 et va être réalisé par le groupe de Laurent Deslandes au LIPME *via* l'expression transitoire d'une construction *pDEX-3HA-turbo-ernA* par rapport à une construction témoin *pDEX-3HA-turbo-NLS-eGFP*. Ce contrôle permet l'adressage de la eGFP dans le noyau tout comme l'est ErnA et nous permet d'éliminer des interactants qui seraient aspécifiques. Cependant en raison de la nécessité de mettre en place le protocole de transformation d'*A. evenia* à Toulouse au préalable qui demande du temps, il a d'abord été effectué cette expérimentation sur *N. benthamiana* qui est réalisée en routine au LIPME. Cette expérimentation sur le tabac est d'autant plus intéres-

[134] Forme anglicisée "d'étiquette", aussi appelée "tag" dans la suite du texte.

A



B

Mediator = multiprotein complex
Recruitment of **RNA polymerase II**

mediator of RNA polymerase II transcription subunit 32-like
mediator of RNA polymerase II transcription subunit 15a isoform X1
mediator of RNA polymerase II transcription subunit 16 isoform X1
mediator of RNA polymerase II transcription subunit 27-like
mediator of RNA polymerase II transcription subunit 15a-like isoform X3
mediator of RNA polymerase II transcription subunit 14
mediator of RNA polymerase II transcription subunit 23 isoform X1
mediator of RNA polymerase II transcription subunit 10b-like isoform X3
mediator of RNA polymerase II transcription subunit 28-like
mediator of RNA polymerase II transcription subunit 12
mediator of RNA polymerase II transcription subunit 14
mediator of RNA polymerase II transcription subunit 13 isoform X1
mediator of RNA polymerase II transcription subunit 22a-like
mediator of RNA polymerase II transcription subunit 6 isoform X2
mediator of RNA polymerase II transcription subunit 17-like
mediator of RNA polymerase II transcription subunit 27

Epigenetic mechanisms
Chromatin remodeling

ATP-dependent helicase BRM (x4)
histone acetyltransferase HAC1-like
histone acetyltransferase HAC1-like isoform X1
chromatin structure-remodeling complex protein SYD
protein MICRORCHIDIA 1-like
SWI/SNF complex subunit SWI3C (x2)
transcription factor GTE8-like isoform X1

Splicing
Cell division cycle

partial, crooked neck-like protein 1
proline-rich protein PRCC
partial, cell division cycle 5-like protein

Transcription factors

transcriptional corepressor LEUNIG-like isoform X3
transcriptional corepressor SEUSS-like isoform X1
topless-related protein 1-like isoform X1
probable transcription-associated protein 1 isoform X1
transformation/transcription domain-associated protein-like isoform
probable inactive poly [ADP-ribose] polymerase SRO3

Other proteins

phytochrome E isoform X1
uncharacterized protein LOC107785315 isoform X2
uncharacterized protein LOC104239499 isoform X1
uncharacterized protein LOC104223121 isoform X1
uncharacterized protein LOC107785315 isoform X1
uncharacterized protein LOC107778684

Figure 69. Identification des interactants d'ErnA par une approche de Turbo-ID chez *Nicotiana benthamiana*.
(**A**) Vulcano plot de l'abondance des peptides identifiés après analyse LC-MS/MS dans la condition 3HA-Turbo-ErnA par rapport à la condition témoin 3HA-turbo-NLS-eGFP. L'encadré vert correspond aux peptides ayant le meilleur score (points jaunes) Les autres points de couleurs indiquent d'autres protéines enrichies dans la condition 3HA-Turbo-ErnA mais avec une significativité plus faible (points oranges) ou des protéines enrichies de la même façon dans la condition témoin et ErnA (points bleus) (**B**) Liste détaillée des 42 protéines identifiées avec le meilleur score et pouvant être des cibles protéiques putatives d'ErnA.

sante à réaliser du fait qu'il a pu être précédemment montré que l'expression transitoire d'*ernA* dans les feuilles induit une HR (figure 67) ; suggérant qu'ErnA doit également impacter fortement la biologie de la plante. De plus, des tests menés sur les différentes formes mutées d'ErnA dans les motifs identifiés dans l'article 1 montrent que la capacité à induire une HR n'est retrouvée que chez les formes actives capables de complémenter le mutant ORS3257Δ*ernA* (figure 67), suggérant que la forme fonctionnelle d'ErnA est à l'origine d'une HR (tests réalisés par Mélanie Carcagno).

Pour l'approche de Turbo-ID, il a ainsi pu être validé la bonne expression des protéines taguées ainsi que leur capacité de biotinylation et d'auto-biotynlation par agro-infiltration de feuilles de tabac après purification sur colonne de streptavidine (figure 68). Enfin, après digestion des protéines biotinylées par la trypsine[135], les peptides ainsi obtenus ont été purifiés et envoyés à analyser en LC MS/MS[136]. L'analyse différentielle de l'abondance des protéines retrouvées dans la condition ErnA par rapport à celle du NLS-eGFP a permis d'identifier 42 protéines candidates présentant un score élevé (figure 69A). Ces candidats peuvent se regrouper en cinq catégories principales : i) un groupe de 16 sous-unités protéiques (23 si on considère la liste complète) composant le complexe Médiateur[137] notamment connu pour recruter l'ARN polymérase II, ii) des facteurs de transcription (6) tels que LEUNIG-like (NbD041060.1), SEUSS-like (NbE44071061.1) ou TOPLESS-related (NbD025179.1), iii) des régulateurs épigénétiques (11) comme des hélicases BRM[138] appartenant au complexe de remodelage de la chromatine SWI/SNF[139] ou encore des histones acétyltransférase HAC1-like[140], iv) des protéines impliquées dans l'épissage des ARNs ou dans le cycle cellulaire (3) et v) d'autres protéines (6) dont la plupart sont de fonction inconnue (figure 69B). Ces résultats préliminaires sont fort prometteurs car ils corroborent les précédentes expérimentations montrant qu'ErnA lie les acides nucléiques (Teulet et al., 2019). En effet, la majorité des catégories protéiques identifiées impliquerait une interaction proche avec de l'ARN ou de l'ADN. Par ailleurs, le nombre élevé de protéines détectées et constituant le complexe Médiateur (16) suggère qu'ErnA pourrait intervenir au niveau de la transcription de gènes cibles pour moduler leur expression. En effet, le Médiateur est un complexe protéique nécessaire à la transcription des gènes par l'ARN polymérase II et ses multiples sous-unités participent à l'expression de gènes spécifiques à travers la reconnaissance de divers facteurs de transcriptions et signaux moléculaires largement décrits chez les mammifères (Yin & Wang, 2014;

[135] Protéase à sérine qui hydrolyse le plus souvent les liaisons peptidiques situées en C-terminal d'un résidu lysine ou arginine.
[136] Liquid Chromatography with tandem Mass Spectrometry. Technique de spectrométrie de masse qui vise à mesurer la masse et la structure chimique de molécules d'intérêts, dans notre cas des fragments peptidiques.
[137] En anglais : Mediator complex. Complexe protéiques retrouvé chez les eucaryotes et dont le nombre de sous-unités peut varier d'une espèce à une autre.
[138] BRAHMA : sous-unité centrale de l'ATPase.
[139] Initialement identifiées chez la levure : mating-type switch (SWI) et sucrose non-fermenting (SNF). Sous-famille de protéines formant un complexe de remodelage de la chromatine dépendant de l'ATP.
[140] Histone acetyltransferase 1-like.

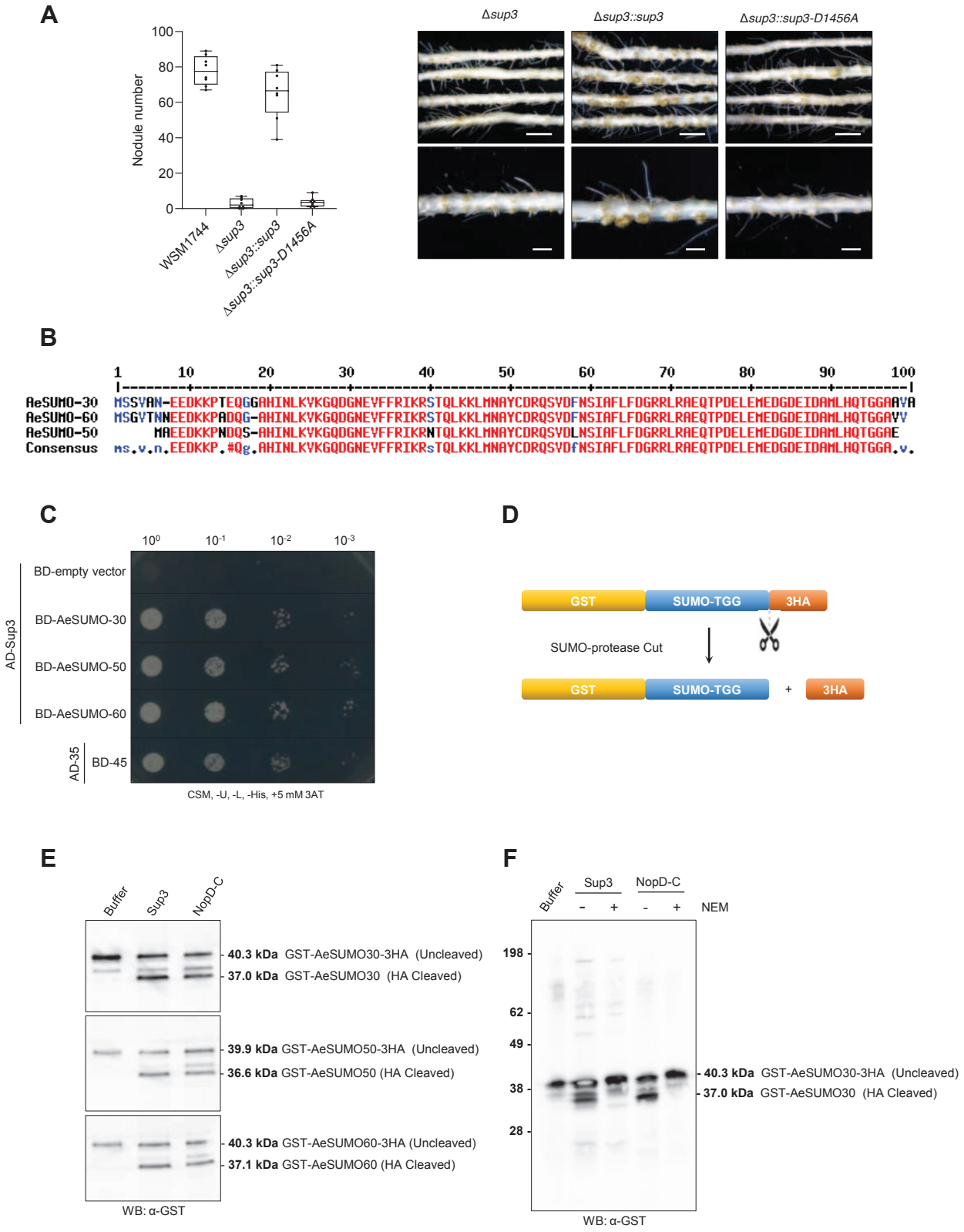


Figure 70. L'activité SUMO-protéase de Sup3 est requise pour sa capacité à déclencher la nodulation.
(**A**) (A gauche) Nombre de nodules sur des racines d'*A. indica* 21 jours après inoculation avec la souche WSM1744 et ses mutants associés. Box plots montrant les résultats d'une expérience réalisée sur 8 plantes. (A droite) Phénotype symbiotique des racines inoculées. Echelle : ligne 1, 0.5 cm; ligne 2, 0.2 cm. (**B**) Alignement des trois SUMO canoniques identifiées chez *A. evenia* et obtenu sur http://multalin.toulouse.inra.fr/multalin/. (**C**) Double-hybride en levure pour tester l'interaction de Sup3 avec les SUMO identifiés chez *A. evenia* (AeSUMO30/AeSUMO50/AeSUMO60). BD-45 et AD-35 sont utilisés comme témoin positif d'une interaction protéine-protéine. Plusieurs dilutions des cultures de levures ont été testées ainsi que du 3AT (3-Amino-1,2,4-triazole) pour inhiber les interactions aspécifiques responsables d'une auto activation du prometteur en absence d'une interaction protéique. (**D**) Principe expérimental du clivage par une SUMO-protéase du tag 3HA. (**E** et **F**) Western-blot anti-GST illustrant l'activité protéase de Sup3. NopD est utilisé comme témoin positif pour le clivage du tag -3HA (Xiang et al., 2020). NopD-C : partie C-terminale de NopD contenant le domaine ULP. En E, l'activité SUMO-protéase de Sup3 et NopD-C a été testée en présence ou pas d'un inhibiteur de protéase NEM : N-ethylmaleimide.

Allen & Taatjes, 2015). Il peut intervenir aussi dans d'autres processus comme par exemple des régulations épigénétiques et la production d'ARNs non-codants (Allen & Taatjes, 2015). Son rôle central dans la transcription fait que le Médiateur est au centre de nombreux processus cellulaires et il a pu être montré chez les plantes qu'il participe entre autres à l'immunité de la plante ainsi qu'à son développement (Yang et al., 2016). La présence de FTs dans les cibles biotinylées suggérerait également l'existence d'un complexe protéique au sein duquel ceux-ci seraient retrouvés avec ErnA pour déclencher la HR observé sur le tabac. Le fait que l'on retrouve par ailleurs des régulateurs épigénétiques en plus de divers FTs dans cette analyse de Turbo-ID laisse suggérer qu'ErnA pourrait avoir un rôle épigénétique en agissant sur des régions spécifiques de l'ADN pour faciliter leur accessibilité et l'expression des gènes. Cependant, l'ensemble de ces résultats préliminaires sont à pondérer car ils ont été obtenus sur un système hétérologue déclenchant une forte réponse immunitaire en réponse à ErnA et qui implique sans doute des acteurs différents que ceux requis pour la nodulation. Des essais sont actuellement en cours pour mener ce même genre d'approche sur *Aeschynomene*. Néanmoins, compte tenu du fort degré de conservation du complexe Médiateur chez les plantes, on peut s'attendre à retrouver ce candidat chez *A. evenia* à la différence des FTs qui pourraient être plus plante-spécifiques dans ce contexte et impliqués dans le processus symbiotique.

2.3. Interaction des ET-Nods avec des partenaires SUMOylés

Dans l'article 1, il a pu être identifié des motifs jouant un rôle important pour l'activité d'ErnA dont notamment un motif SIM[141] putatif dont le mutant associé perd son activité de nodulation. Le domaine SIM est connu pour permettre une interaction non-covalente entre un peptide SUMO et la protéine possédant un tel motif (Lascorz et al., 2022). Ce type d'interaction peut conduire à une variété de modifications de la protéine SUMOylée, allant de sa structure à sa fonction biologique (Lascorz et al., 2022). En parallèle dans cette même étude, il a pu être identifié un nouvel ET-Nod, Sup3, qui tout comme Bel2-5 code pour une putative SUMO-protéase responsable de la déSUMOylation de protéines cibles. Cependant l'ensemble de ces observations n'a pas fait l'objet de validations fonctionnelles et il nous est donc apparu nécessaire de confirmer qu'un lien entre SUMOylation et ET-Nod pouvait exister.

C'est pourquoi il a d'abord été construit un mutant de Sup3 dans lequel un des acides-aminés de la triade catalytique du site actif de la SUMO-protéase a été modifié, afin de déterminer si cette activité enzymatique intervient dans l'induction de la nodulation. Des tests effectués sur *A. indica* ont permis de montrer que ce mutant est incapable de complémenter un mutant de délétion WSM1744Δ*sup3* à l'inverse de la forme sauvage de Sup3 (figure 70A). L'activité d'induction de la nodulation de Sup3 est

[141] SUMO-interacting motif.

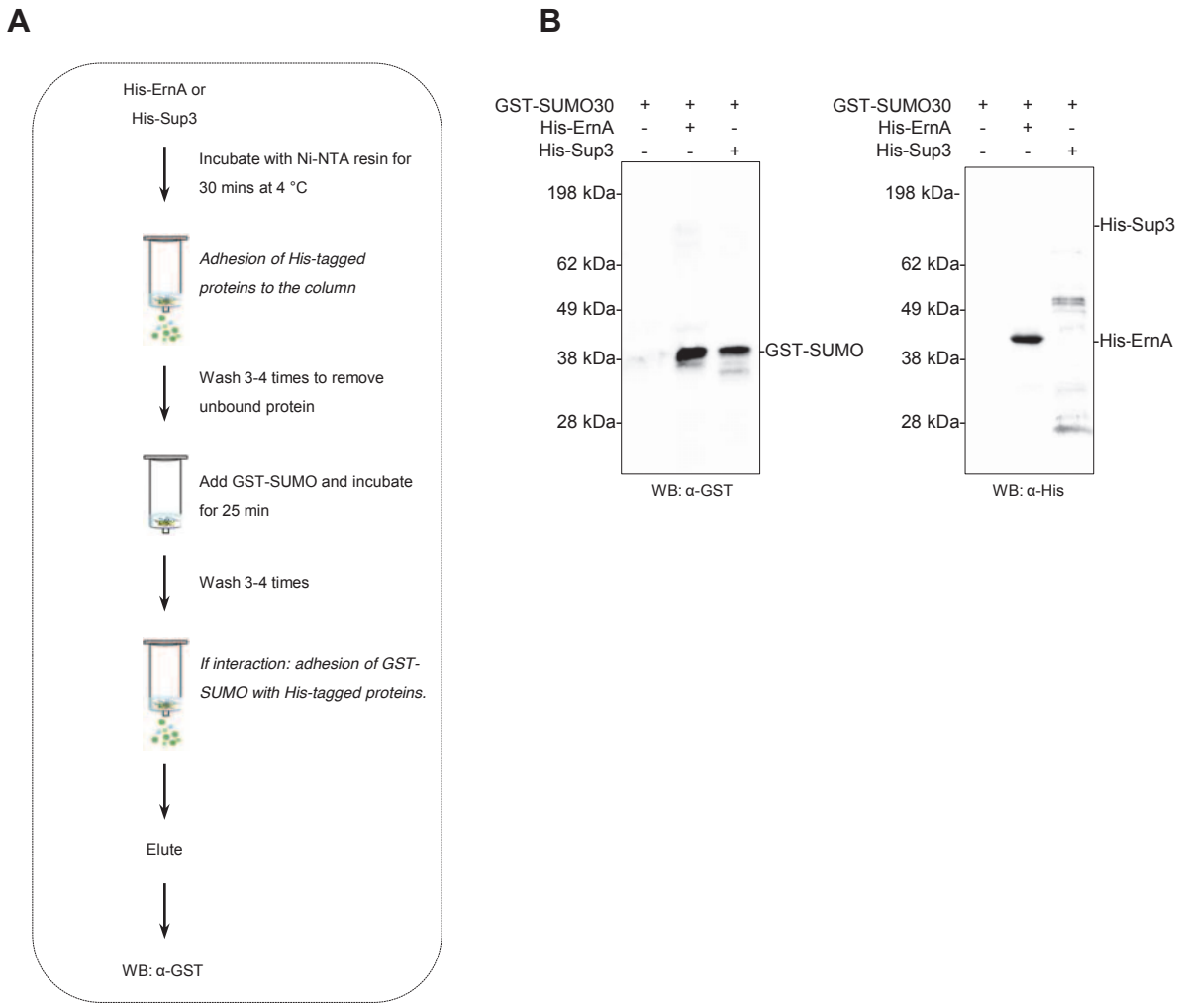


Figure 71. ErnA interagit avec un peptide SUMO.
(**A**) Principe de la Co-IP réalisée pour mettre en évidence une interaction potentielle entre ErnA et les peptides SUMO. Ni-NTA : microbilles magnétiques (nickel) adhérant à la colonne et utilisées pour purifier des protéines tagguées His. (**B**) Western-blot anti-GST (à gauche) illustrant l'interaction entre ErnA et le peptide AeSUMO30 ; et anti-His (à droite) validant la présence des protéines tagguées. Sup3 est utilisé en contrôle positif, la protéine purifiée apparaît fortement dégradée (multiples bandes).

donc à relier avec sa capacité à déSUMOyler certaines cibles de l'hôte. Cette même observation a pu être faite chez des mutants dans le site catalytique du domaine SUMO-protéase de Bel2-5, qui présentent également une altération de la nodulation (Ratu, Hirata, et al., 2021).

Parallèlement, il a été testé par le groupe de Peter Mergaert à l'I2BC la capacité de Sup3 ou d'ErnA à interagir avec les trois peptides SUMO canoniques retrouvés chez *A. evenia* (figure 70B). L'interaction entre Sup3 et les SUMO a d'abord été confirmée sur un système hétérologue par double-hybride en levure. Comme attendue de par sa fonction prédite, Sup3 est capable d'interagir avec les trois SUMO testées puisqu'une croissance des levures est observée sur un milieu sélectif (figure 70C), soulignant la bonne activation du gène rapporteur associée à une interaction entre nos deux protéines. La capacité de déSUMOylation de Sup3 a pu être aussi vérifiée *in vitro* comme décrit par Xiang et ses collaborateurs (Xiang et al., 2020) à travers le clivage d'un motif 3HA fusionné en C-terminal de la protéine (figure 70D et 70E). L'utilisation d'un inhibiteur de protéase (NEM[142]) a permis de vérifier que la bande de plus faible poids moléculaire observée en présence de Sup3 est bien due à son activité protéolytique (figure 70F). Enfin, la capacité d'ErnA à interagir avec les SUMOs à travers son motif SIM a été validée par Co-IP d'une forme tagguée 6His d'ErnA en N-terminal avec un peptide SUMO fusionné à la GST (figure 71).

L'ensemble de ses résultats préliminaires suggère une interaction possible des ET-Nods avec des partenaires SUMOylés mais à travers différents modes d'interaction. Ces résultats sont bien-sûr à confirmer et des mutants déjà disponibles (tels que ceux dans le motif SIM d'ErnA ou dans le domaine SUMO-protéase de Sup3) seront ajoutés comme contrôle de ces tests *in vitro*.

A noter que les matériels et méthodes utilisées dans ce chapitre sont disponibles en annexe 2.

[142] N-ethylmaleimide.

Chapitre 4

1. Contexte de l'étude

La collecte des souches de *Bradyrhizobium* auprès de différents laboratoires éparpillés dans le monde et qui ont servi pour l'étude du chapitre 1 nous a amené à collaborer avec le groupe d'Haiwei Luo basé à Hong-Kong. A l'origine d'un séquençage massif de nouveaux génomes de *Bradyrhizobium* issus principalement du riz (*Oryza sativa indica*), ces nouvelles données ont permi d'agrandir considérablement la diversité phylogénomique de ce genre bactérien. Il s'est notamment avéré que parmi les nouvelles souches identifiées, 208 se retrouvent positionnées au sein du supergroupe des *Bradyrhizobium* photosynthétiques. Ces données génomiques constituent une nouvelle ressource très importante pour mieux comprendre la biologie de ce groupe phylogénétique aux propriétés bien particulières pour lequel 13 souches avaient été jusqu'alors séquencées – la plupart étant isolées de nodules d'*Aeschynomene* spp.

L'analyse de ces nouveaux génomes a permis de montrer l'existence de trois clades au sein de ce supergroupe. Le clade 1, qui était jusqu'à présent le principal identifié, apparaît maintenant lui-même constitué de 6 sous-clades ou sous-populations. Une caractéristique commune de l'ensemble de ces nouvelles souches séquencées et qu'elles possèdent toutes les gènes *nifHDK* nécessaires à la synthèse du complexe enzymatique de la nitrogénase ainsi que le gène *nifV* codant pour une homocitrate synthase impliquée dans la synthèse du FeMo-co le plus souvent absent chez les rhizobiums. Ceci suggère que la fixation d'azote à l'état libre serait possible pour l'ensemble de ces souches ; ce qui a pu être vérifié expérimentalement pour un certain nombre d'entre-elles. Un autre point commun retrouvé chez ces souches est l'absence des gènes canoniques *nodABC*. Alors que le caractère photosynthétique était considéré comme un trait phénotypique majeur de ce phylogroupe, il est apparu que les gènes photosynthétiques sont absents chez plusieurs sous-populations caractérisées.

En parallèle de cette analyse génomique, nous avons examiné les propriétés symbiotiques d'une sélection de souches représentatives de la nouvelle diversité existante. Bien qu'aucune des nouvelles souches testées n'aient initialement été isolées de nodules, toutes ont été montrées capables de noduler *A. indica*, plus ou moins efficacement. Cette capacité universelle à noduler *A. indica* suggère que toutes les souches du supergroupe des *Bradyrhizobium* photosynthétiques possèdent les déterminants symbiotiques gouvernant la symbiose Nod-indépendante, T3SS-indépendante et que ces derniers pourraient être essentiels à leur survie. Ceci pourrait alors expliquer qu'aucun mutant de souches de *Bradyrhizobium* photosynthétiques ayant perdu complètement sa capacité de nodulation n'ait été jusqu'alors sélectionné (Bonaldi et al., 2010). Dès lors, cette étude offre une diversité génétique nouvelle et des ressources génomiques importantes qui pourraient nous aider à révéler les

déterminants gouvernant la symbiose Nod-indépendante, T3SS-indépendante par de nouvelles approches (gain de fonction) qui n'ont pas encore été utilisées à ce jour (cf. discussion et perspectives).

2. Article 3 : "Phylogenetic and population structure of the *nod*-free but nodulating *Bradyrhizobium* phylogroup"

Les données supplémentaires (Dataset S1 à S4) associées à cette étude sont disponibles sur : https://www.biorxiv.org/content/10.1101/2023.06.14.544914v2.supplementary-material

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Phylogenetic and population structure of the *nod*-free but nodulating *Bradyrhizobium* phylogroup

Lu Ling[1,2^], Alicia Camuel[3,4^], Sishuo Wang[1], Xiaojun Wang[2], Xinqin Lin[2], Nico Nouwen[3,4], Eric Giraud[3,4*], Haiwei Luo[1,5*]

[1]Simon F. S. Li Marine Science Laboratory, School of Life Sciences and State Key Laboratory of Agrobiotechnology, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

[2]Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China

[3]IRD, Laboratoire des Symbioses Tropicales et Méditerranéennes (LSTM), UMR IRD/ SupAgro/INRA/ Université de Montpellier /CIRAD, TA-A82/J– Campus de Baillarguet 34398 Montpellier cedex 5, France

[4]PHIM Plant Health Institute, Université Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France

[5]Institute of Environment, Energy and Sustainability, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

[^]**These are co-first authors**: Lu Ling, Alicia Camuel

[*]**These are co-corresponding authors:** Eric Giraud (eric.giraud@ird.fr), Haiwei Luo (haiweiluo@cuhk.edu.hk)

Key words: *Bradyrhizobium*, *Aeschynomene*, nodulation, population genomics

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Bradyrhizobium is a main rhizobial lineage of which the majority of its members nodulate legume plants using Nod factors (NFs) encoded by the *nod* genes. However, the "photosynthetic" phylogroup within *Bradyrhizobium* (PB) does not contain *nod* genes but are still capable of establishing a nitrogen-fixing symbiosis with some tropical legumes of the *Aeschynomene* genus. Furthermore, members of this phylogroup colonize rice roots and promote plant growth. Despite this unique combination of ecological features, genomes of only 13 PB members are sequenced, and almost all are isolated from *Aeschynomene* spp. nodules. Here, we sequenced 208 new PB members predominantly associated with rice. The extended PB phylogroup comprises three main clades: a basal clade with significant expansion of its diversity, followed by the diversification of a singleton clade from a new clade exclusively represented by our strains. Although the PB strains universally lack the canonical *nod* genes, all 27 assayed strains representing the broad diversity of these clades induced nodules on *Aeschynomene indica*. Interestingly, significant differences in the efficiency of symbiosis between the main clades were observed and the singleton clade showed an intermediate symbiotic phenotype aligning well with its intermediate phylogenetic position. Our strain collection expands the phylogenetic and ecological diversity of the *nod*-free but nodulating *Bradyrhizobium* and shows that the NF-independent nodulation of *Aeschynomene* is a common trait of this phylogroup, in contrast to the photosynthetic trait originally thought as its unifying feature.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Bradyrhizobium is one of the largest and most diverse rhizobial genera and the primary symbiont of a wide range of legumes [1, 2]. The genus *Bradyrhizobium* encompasses seven phylogenetic supergroups (phylogroups) [3, 4], among which the "photosynthetic" *Bradyrhizobium* (PB) supergroup is one of the most special. First, these bacteria are photosynthetic which is a very rare property in rhizobia [3]. Second, members belonging to this supergroup can form nitrogen-fixing nodules not only on roots but also on stems of some tropical semi-aquatic species of the *Aeschynomene* genus [5]. Finally, the most original feature of the PB strains is the fact that these nodulating bacteria lack the canonical *nodABC* genes encoding the enzymes that synthesize the core structure of the nodulation factors (NF).

The available evidence indicates that the PB phylogroup takes a novel symbiotic strategy that differs from the traditional universal NF-based pathway [6, 7]. Although several attempts have been made to identify the symbiotic determinants of PB strains by mutagenesis, the mechanism has not yet been discovered [8]. This mystery has become more intriguing since the recent discovery of another NF-independent symbiotic process present in a wide variety of non-photosynthetic *Bradyrhizobium* strains [9]. This last process relies on a Type III secretion system (T3SS) which secretes, into the plant cells, effectors (T3Es) able of triggering nodulation [10-12]. The absence of T3SS in most PB strains indicates that they use a distinct symbiotic mechanism that was shown specific to this phylogroup [12].

Despite these unique features, the research on photosynthetic *Bradyrhizobium* has been very limited. Currently, there are only 13 PB strains that have their genomes sequenced, largely limiting comparative genomics analysis on this important *Bradyrhizobium* phylogroup. Although most PB strains were isolated from *Aeschynomene* spp. nodules, the PB phylogroup was also

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

detected in paddy soil, rice roots as well as in lake water [13-16], suggesting that its members have much wider ecological niches than previously thought. Here, we cultured and sequenced the genomes of 208 isolates (Fig. 1) predominantly associated with three niches of rice (within root, rhizosphere, bulk soil). We then conducted genomic analyses and physiological assays to reveal the phylogenetic and population structure of the photosynthetic *Bradyrhizobium* supergroup.

Results and discussion

The new genomes expand the phylogenetic diversity of the photosynthetic supergroup of *Bradyrhizobium*

With the 208 newly isolated members (see Supplementary Text for photosynthetic *Bradyrhizobium* isolation), the phylogeny of the PB supergroup is split into three deep-branching clades (see Fig. S1 for their phylogenetic position in the species tree of the entire *Bradyrhizobium* genus). Clade 1 represents the evolutionarily basal lineage. It initially comprised only 12 strains, but it now has its diversity expanded by having 68 new strains representing several new lineages delineated as distinct genetically isolated populations (Fig. 1, Datasets S1 and S2; also discussed in the next section). Clade 2 is a singleton clade with only one strain (*Bradyrhizobium* sp. STM3843) available and represents a phylogenetic intermediate among the three clades. Clade 3 is new and consists exclusively of 140 new strains. The evolutionary branching order of the three clades is verified with two outgroup-independent methods (Fig. 1, Fig. S7A) and the outgroup-dependent method (Fig. S1). Collectively, our new strains appreciably increase the existing phylogenetic diversity of the PB supergroup. Comparative genomic analysis shows that except for *Bradyrhizobium* sp. ORS285, which is known to have

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

93 *nod* genes and use both a *nod*-dependent and -independent strategy for nodulation [17, 18], all
94 PB supergroup members lack the *nod* genes.
95
96 Fine-scale population structure of the photosynthetic supergroup of *Bradyrhizobium*
97 Apart from the broad diversity of the PB supergroup, we further asked whether fine-scale
98 phylogenetic differentiation correlates with the ecological niches where the PB supergroup
99 members were found. Of the 208 new PB supergroup strains, 205 were collected from three rice
100 plants located in the same rice field with each plant 5-10 m apart from the others. The remaining
101 three strains were isolated from grassland and forest. Bacterial isolation was performed from
102 within the root, rhizosphere, and bulk soil for each plant, collectively giving rise to 145, 23 and
103 40 PB strains, respectively (Dataset S1). To facilitate the correlative analysis between fine-scale
104 phylogenetic groups and the niches, we applied the population genomics software PopCOGenT
105 [19], which defines the boundaries of populations (reported as "main cluster" or MC) between
106 which gene flow is limited, distinguishes recent gene flow from historical events, and detects
107 subpopulations (reported as 'subclusters') under ongoing differentiation within a population.
108 Using this approach, we showed that all strains in Clade 3 share the membership of a
109 single population (MC1), whereas strains from Clade 1 fall into many populations. In Clade 1,
110 MC2, MC3, MC4, MC5 and MC6 are each exclusively comprised of the new strains, whereas
111 MC7 and the remaining unassigned isolates are from published resources. In Clade 1, the within-
112 and between-population similarity was above and below 95% ANI (Fig. S2), respectively,
113 consistent with the operational species threshold of 95% ANI [20], whereas the 16S rRNA gene
114 shows little divergence with the between-population similarity generally above 99% (Fig. S2,
115 Dataset S3). Based on the available samples, intra-population subdivision occurred within MC1,

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

MC2 and MC4. We found that each population and subpopulation have members sampled from different plants and niches (Fig. 1), suggesting that the PB populations sampled from the rice field are not separated or subdivided according to the physical separation between plant individuals and the three typical niches (root, rhizosphere, and bulk soil).

Rather, we found an interesting association between important metabolic pathways and population identity. Remarkably, although the photosynthetic supergroup was named by the presence of the photosynthesis genes [3], these genes are exclusively and universally found in MC2, MC3, MC6 and MC7 but completely missing from MC1, MC4 and MC5 (Figure 1, dataset S4) as well as STM3843 as previously reported [21, 22]. Our result indicates that photosynthesis is far from being a common characteristic trait that defines the PB phylogroup.

Also interesting is the prevalence of a unique Type III secretion system (T3SS) in MC2, MC4 and MC5 but completely missing in other populations (Fig. 1, Fig. S3B and S3C, Dataset S4). It is one of the six T3SS subtypes identified from known *Bradyrhizobium* phylotypes [23]. It was previously identified only in the PB strain *Bradyrhizobium oligotrophicum* S58 (thus named 'S58-T3SS' subtype) and its function remains unknown [23]. Since the main function of T3SS in rhizobia is to translocate effector proteins into host cells that modulate the host immune response or trigger nodulation [24], it cannot be excluded that this T3SS type specifically identified in PB members plays a role during their interaction with their host plants (*Aeschynomene* spp. and rice).

We also found that *mxa* gene sets participating in methanol metabolism were specific to MC1 (Dataset S4), which suggests that these MC1 strains may form intimate associations with plants considering that methanol is a waste product formed during plant-cell wall degradation [25]. Our methanol metabolism experiment showed that MC1 strains could not use methanol as a

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

sole C source (Fig. S4), hinting that the expression of *mxa* genes in these MC1 strains might be under regulation of unknown mechanisms.

The ability to nodulate *Aeschynomene indica* is a conserved trait shared by all PBs but differs between the three major PB clades

Since the new PB strains were isolated from non-legume environments, we thought that it would be very interesting to investigate (i) whether they are capable of performing nitrogen fixation under free-living and symbiotic states, (ii) whether they are capable of nodulating *Aeschynomene indica*, the typical host plant of the previously isolated PB strains, and (iii) whether there is a difference in nodulation ability across the clades and the populations. The detailed assay procedures can be seen in Supplementary Text.

All PB supergroup members have a *nif* gene cluster necessary for nitrogenase fixation. Interestingly, duplication of the *nifH* gene which encodes one of the structural proteins of the nitrogenase enzyme complex was observed in all PB strains (Fig. S3A), although the functional consequence is not known. Across the PB supergroup members, an important difference was observed regarding how the *nif* genes are structured. In most populations, all *nif* genes are co-located, along with other genes potentially involved in nitrogen fixation. This is not the case for a few members such as some basal lineages (e.g., *Bradyrhizobium* sp LMG 8443) and some members assigned to several populations (e.g., SZCCHNR1015 in MC3 and *Bradyrhizobium* sp. 83002 in MC7), where the *nif* genes were found in two contigs. This is likely due to structural rearrangement, but a DNA assembly artefact cannot be ruled out.

Another original feature observed of the PB *nif* cluster is the universal presence of the *nifV* gene encoding for a homocitrate synthase that is involved in the biosynthesis of the

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

nitrogenase cofactor (FeMo-Co), and which is absent in most other rhizobial lineages [26, 27]. An acetylene reduction assay on 27 representative PB strains covering all three major clades, all newly identified populations (MC1 to MC6) within Clade 1 and Clade 3, and most of the identified subpopulations within each population, confirmed their ability to fix nitrogen under free-living conditions (Fig. S5A), a trait absent in most rhizobia [27].

The plant inoculation experiments showed that all 27 representative strains were also able to induce nodules on *A. indica* (Fig. 2B, 2C). However, a significant difference in symbiosis efficiency was observed between the major clades. Clade 1 strains had the same symbiotic properties as the model strain ORS278, inducing a large number of nitrogen-fixing nodules that stimulated the growth of the plants (Fig. 2A, 2B, Fig. S5B and Fig. S6A). In contrast, Clade 3 strains formed a much smaller number of nodules compared to Clade 1 strains and the measured nitrogenase enzyme activity was lower under the symbiotic condition (Fig. 2B, Fig. S5B). Microscopic analysis showed that the nodules elicited by Clade 3 strains displayed multiple altered phenotypes: i) some nodules contained necrotic areas, ii) in others the central tissue was completely digested, and finally, iii) the nodules that displayed less drastic symptoms contained mainly dead bacteria (Fig. 2C and 2F). These observations indicate that besides inducing fewer nodules, Clade 3 strains are unable to maintain an effective chronic infection, which explains why they have no beneficial effect on plant growth (Fig. 2A, 2C and 2F). Interestingly, the singleton Clade 2 represented by strain STM3843 which is phylogenetically positioned between Clade 1 and Clade 3, showed an intermediate phenotype. The number and phenotype of the nodules is comparable to those of Clade 1 strains but stimulation of plant growth is less (Fig. 2A, 2B and 2E).

The symbiotic differences observed between the major clades may result from an

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

inability of the Clade 3 strains to cope with the plant's immune response, and/or an over-induction of the plant's defense mechanisms due to the absence or not properly recognized symbiotic signals. However, it is remarkable that all strains from the PB phylogroup, which come from different origins and geographical locations, are capable of nodulating *A. indica*.

Concluding remarks

The present study greatly expands the phylogenetic and ecological diversity of the "photosynthetic" *Bradyrhizobium* supergroup. The photosynthetic trait, which was thought to be a unifying character of these bacteria, in fact displays a patchy distribution within the supergroup. In contrast, the ability to develop a NF-independent symbiosis with *Aeschynomene* spp., despite varying in nodulation efficiency between major clades, appears to be a universal trait of the phylogroup. We therefore propose to rename this phylogroup as "*nod*-free but nodulating *Bradyrhizobium"*.

The conservation of the nodulation character in such an ecologically diverse phylogroup and the fact that these strains were mainly isolated from the rice environment lacking selective pressure from legumes suggest that the symbiotic determinants governing the NF-independent, T3SS-independent symbiosis may be also essential to their life. This could explain the failures so far to identify these symbiotic determinants using loss-of-function approaches. Although it is beyond our goal to reveal these symbiotic determinants in the present study, the resulting new knowledge leads us to consider tackling this difficult problem using other methods (e.g., gain-of-function approaches). Furthermore, these data lay the ground for a new evolutionary concept of an alternative symbiotic process between rhizobia and legumes in which key symbiotic genes can be inherited vertically rather than transferred horizontally as described for *nod* genes.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In addition to their ability to use a unique symbiotic process with legumes, PB strains may also have the original function of colonizing rice roots and improving plant growth [17, 18, 28]. This new wealth of genomic sequences of strains isolated from rice could therefore be of great value in the future to improve our understanding of this plant growth stimulating effect.

Acknowledgements

We thank Xiaoyuan Feng for helping with genome assembly and Jinjin Tao for discussion. This work was supported by the Hong Kong Research Grants Council Area of Excellence Scheme (AoE/M-403/16) and a grant from the French National Research Agency ("ET-Nod"; ANR-20-CE20-0012).

Data availability

The genomic sequences and raw reads of the 208 newly sequenced *Bradyrhizobium* isolates are in the process of uploading to NCBI (Project ID: PRJNA983111) and will be released in the next version of the manuscript.

Competing interests

The authors declare no competing interests in relation to this work.

References

1. Parker MA. The spread of Bradyrhizobium lineages across host legume clades: from Abarema to Zygia. *Microb Ecol* 2015; **69:** 630-640.

2. Sprent JI, Ardley J, James EK. Biogeography of nodulated legumes and their nitrogen-fixing

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

symbionts. *New Phytol* 2017; **215:** 40-56.

3. Avontuur JR, Palmer M, Beukes CW, Chan WY, Coetzee MPA, Blom J, et al. Genome-informed Bradyrhizobium taxonomy: where to from here? *Syst Appl Microbiol* 2019; **42:** 427-439.

4. Ormeno-Orrillo E, Martinez-Romero E. A Genomotaxonomy View of the Bradyrhizobium Genus. *Front Microbiol* 2019; **10:** 1334.

5. Giraud E, Fleischman D. Nitrogen-fixing symbiosis between photosynthetic bacteria and legumes. *Photosynthesis Research* 2004; **82:** 115-130.

6. Lerouge P, Roche P, Faucher C, Maillet F, Truchet G, Prome JC, et al. Symbiotic host-specificity of Rhizobium meliloti is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* 1990; **344:** 781-784.

7. Oldroyd GED. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology* 2013; **11:** 252-263.

8. Bonaldi K, Gourion B, Fardoux J, Hannibal L, Cartieaux F, Boursot M, et al. 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* 2010; **23:** 760-770.

9. Okazaki S, Tittabutr P, Teulet A, Thouin J, Fardoux J, Chaintreuil C, et al. Rhizobium-legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. *ISME J* 2016; **10:** 64-74.

10. Teulet A, Busset N, Fardoux J, Gully D, Chaintreuil C, Cartieaux F, et al. The rhizobial type III effector ErnA confers the ability to form nodules in legumes. *Proc Natl Acad Sci U S A* 2019; **116:** 21758-21768.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

11. Ratu STN, Teulet A, Miwa H, Masuda S, Nguyen HP, Yasuda M, et al. Rhizobia use a pathogenic-like effector to hijack leguminous nodulation signalling. *Sci Rep* 2021; **11:** 2034.

12. Camuel A, Teulet A, Carcagno M, Haq F, Pacquit V, Gully D, et al. Widespread Bradyrhizobium distribution of diverse Type III effectors that trigger legume nodulation in the absence of Nod factor. *ISME J* 2023; **In press**.

13. Chaintreuil C, Giraud E, Prin Y, Lorquin J, Ba A, Gillis M, et al. Photosynthetic bradyrhizobia are natural endophytes of the African wild rice Oryza breviligulata. *Appl Environ Microbiol* 2000; **66:** 5437-5447.

14. Piromyou P, Greetatorn T, Teamtisong K, Okubo T, Shinoda R, Nuntakij A, et al. Preferential association of endophytic bradyrhizobia with different rice cultivars and its implications for rice endophyte evolution. *Appl Environ Microbiol* 2015; **81:** 3049-3061.

15. Van Berkum P, Eardly BD. The aquatic budding bacterium Blastobacter denitrificans is a nitrogen-fixing symbiont of Aeschynomene indica. *Appl Environ Microbiol* 2002; **68:** 1132-1136.

16. Okubo T, Fukushima S, Itakura M, Oshima K, Longtonglang A, Teaumroong N, et al. Genome analysis suggests that the soil oligotrophic bacterium Agromonas oligotrophica (Bradyrhizobium oligotrophicum) is a nitrogen-fixing symbiont of Aeschynomene indica. *Appl Environ Microbiol* 2013; **79:** 2542-2551.

17. Bonaldi K, Gargani D, Prin Y, Fardoux J, Gully D, Nouwen N, et al. 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* 2011; **24:** 1359-1371.

18. Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E, Avarre JC, et al. Legumes symbioses:

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

277 absence of Nod genes in photosynthetic Bradyrhizobia. *Science* 2007; **316:** 1307-1312.

278 19. Arevalo P, VanInsberghe D, Elsherbini J, Gore J, Polz MF. A Reverse Ecology Approach
279 Based on a Biological Definition of Microbial Populations. *Cell* 2019; **178:** 820-834 e814.

280 20. Konstantinidis KT, Tiedje JM. Genomic insights that advance the species definition for
281 prokaryotes. *Proc Natl Acad Sci U S A* 2005; **102:** 2567-2572.

282 21. Mornico D, Miche L, Bena G, Nouwen N, Vermeglio A, Vallenet D, et al. Comparative
283 genomics of aeschynomene symbionts: insights into the ecological lifestyle of nod-independent
284 photosynthetic bradyrhizobia. *Genes (Basel)* 2011; **3:** 35-61.

285 22. Miché L, Moulin L, Chaintreuil C, Contreras-Jimenez JL, Munive-Hernández JA, Del
286 Carmen Villegas-Hernandez M, et al. Diversity analyses of Aeschynomene symbionts in Tropical
287 Africa and Central America reveal that nod-independent stem nodulation is not restricted to
288 photosynthetic bradyrhizobia. *Environ Microbiol* 2010; **12:** 2152-2164.

289 23. Teulet A, Gully D, Rouy Z, Camuel A, Koebnik R, Giraud E, et al. Phylogenetic distribution
290 and evolutionary dynamics of nod and T3SS genes in the genus Bradyrhizobium. *Microb Genom*
291 2020; **6**.

292 24. Teulet A, Camuel A, Perret X, Giraud E. The Versatile Roles of Type III Secretion Systems
293 in Rhizobium-Legume Symbioses. *Annu Rev Microbiol* 2022; **76:** 45-65.

294 25. Nemecek-Marshall M, MacDonald RC, Franzen JJ, Wojciechowski CL, Fall R. Methanol
295 Emission from Leaves (Enzymatic Detection of Gas-Phase Methanol and Relation of Methanol
296 Fluxes to Stomatal Conductance and Leaf Development). *Plant Physiol* 1995; **108:** 1359-1368.

297 26. Hakoyama T, Niimi K, Watanabe H, Tabata R, Matsubara J, Sato S, et al. Host plant genome
298 overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. *Nature* 2009; **462:** 514-
299 517.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

300 27. Nouwen N, Arrighi J-F, Cartieaux F, Chaintreuil C, Gully D, Klopp C, et al. The role of
301 rhizobial (NifV) and plant (FEN1) homocitrate synthases in Aeschynomene/photosynthetic
302 Bradyrhizobium symbiosis. *Scientific Reports* 2017; **7:** 448.
303 28. Cai W, Ou F, Staehelin C, Dai W. Bradyrhizobium sp. strain ORS278 promotes rice growth
304 and its quorum sensing system is required for optimal root colonization. *Environ Microbiol Rep*
305 2020; **12:** 656-666.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

306 **Figure legends**

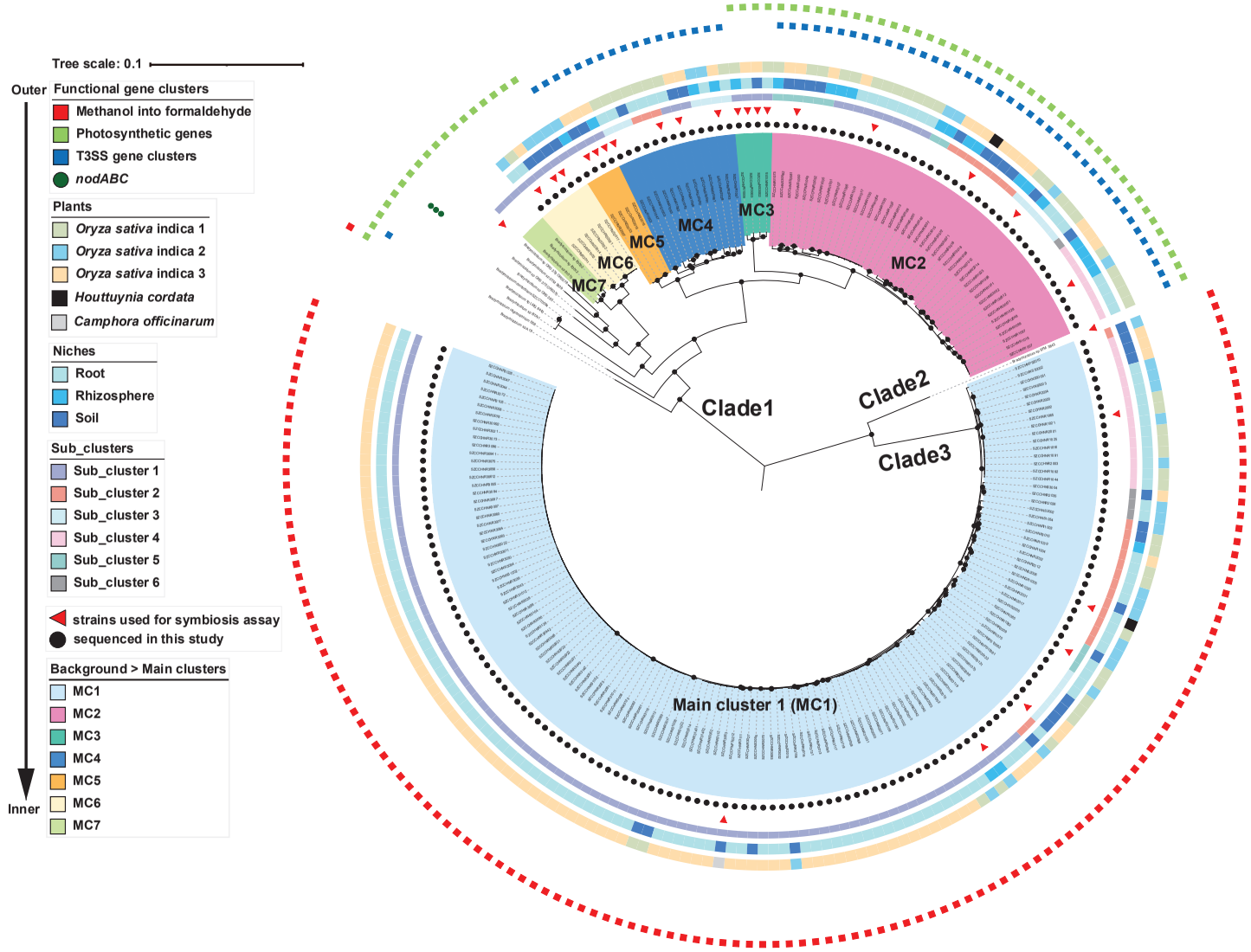
307 **Figure 1.** The phylogenomic tree and population delineation of the "photosynthetic"
308 *Bradyrhizobium* (PB) supergroup. The phylogenomic tree was rooted using the minimum
309 variance (MV) method. The 208 PB strains sequenced in the present study are indicated by black
310 dots in the innermost layer surrounding the phylogeny. The ultrafast bootstrap values higher than
311 or equal to 95% calculated by IQ-Tree were labeled on the nodes with black circles.

312

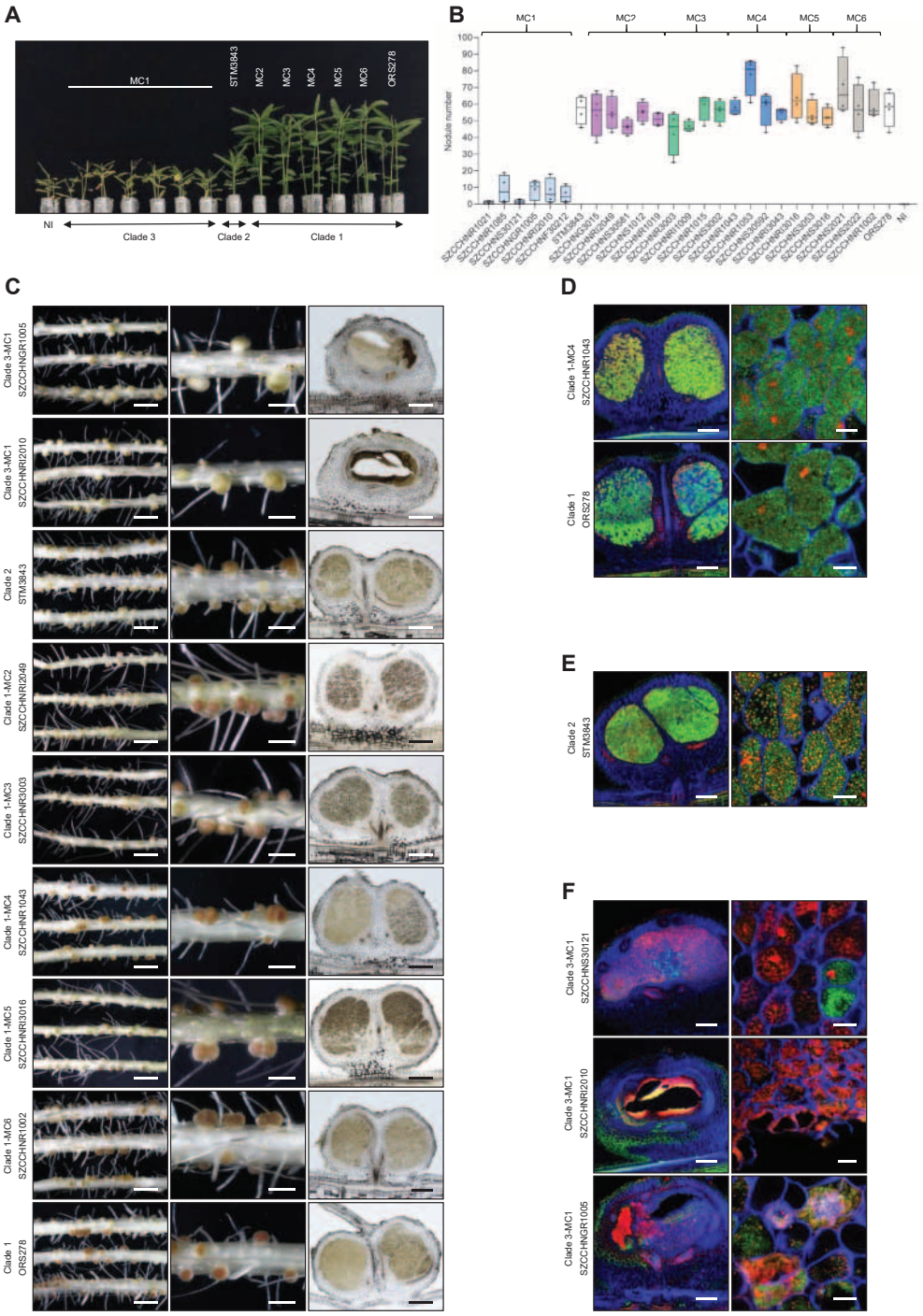
313 **Figure 2.** Symbiotic properties of several representative strains of the main populations
314 identified in the PB supergroup. (A) Comparison of growth of the *A. indica* plants (leaf
315 phenotype) non-inoculated (NI) or inoculated with different representative strains of PB. The six
316 MC1 representative strains are presented in the following order: SZCCHNR1021,
317 SZCCHNR1085, SZCCHNS30121, SZCCHNGR1005, SZCCHNRI2010 and SZCCHNF30212.
318 For the other MCs, only one representative strain is shown: SZCCHNRI2049 (MC2),
319 SZCCHNS3002 (MC3), SZCCHNS30592 (MC4), SZCCHNRI3016 (MC5) and SZCCHNS2021
320 (MC6), Photos were taken 17 days after inoculation. (B) Nodule number on *A. indica* plants
321 induced by different representative strains of PB at 17 days after inoculation. Box plots show the
322 results of one of the two experiments performed independently (4 plants each). (C) Symbiotic
323 phenotypes of some representative strains tested on *A. indica*. Column 1 and 2: photo of roots
324 and nodules. Scale bars: column 1, 0.5 cm; column 2, 0.2 cm. Column 3: Micro-sections of
325 nodules observed using light microscopy. Scale bars: 250 μm. (D) Confocal microscopy images
326 of micro-sections of nodules elicited by Clade 1 strains. The nodules formed by
327 SZCCHNRI3016 from MC4 and ORS278 are shown as examples. The central nodule tissue is
328 intracellularly infected with live spherical bacteroids stained green by SYTO9. (E) Confocal

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

329 microscopy images of micro-section of nodules elicited by Clade 2 strain STM3843. The nodules
330 look normal, but a mix of live and dead bacteria (stained red by propidium iodide) can be
331 noticed. (F) Confocal microscopy images of micro-section of nodules elicited by Clade 3 strains
332 displaying different phenotypes: nodule containing mainly dead bacteria (SZCCHNS30121),
333 nodule with a completely digested central tissue (SZCCHNRI2010); and nodule with a necrotic
334 area and a digested zone (SZCCHNGR1005). Scale bars for (D), (E) and (F): column 1, 200 μm;
335 column 2, 10 μm.
336


Tree scale: 0.1
Outer
Inner
Functional gene clusters
Methanol into formaldehyde
Photosynthetic genes
T3SS gene clusters
nodABC
Plants
Oryza sativa indica 1
Oryza sativa indica 2
Oryza sativa indica 3
Houttuynia cordata
Camphora officinarum
Niches
Root
Rhizosphere
Soil
Sub_clusters
Sub_cluster 1
Sub_cluster 2
Sub_cluster 3
Sub_cluster 4
Sub_cluster 5
Sub_cluster 6
strains used for symbiosis assay
sequenced in this study
Background > Main clusters
MC1
MC2
MC3
MC4
MC5
MC6
MC7
MC7
MC6
MC5
MC4
MC3
MC2
Clade1
Clade2
Clade3
Main cluster 1 (MC1)

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.14.544914; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Supplementary Information

Phylogenetic and population structure of the *nod*-free but nodulating *Bradyrhizobium* phylogroup

Lu Ling[^], Alicia Camuel[^], Sishuo Wang, Xiaojun Wang, Xinqin Lin, Nico Nouwen, Eric Giraud[*], Haiwei Luo[*]

[*]**These are co-corresponding authors:** Eric Giraud (eric.giraud@ird.fr), Haiwei Luo (haiweiluo@cuhk.edu.hk)

This PDF file includes:

Supplementary Text. Methods

Supplementary references.

Figures S1 to S7.

Supplementary Text: Methods

Plants and soils collection

Three rice plants (*Oryza sativa* indica) were collected from the same paddy field in Hunan province, China (27.948 °N, 113.221 °E) on 12 July 2021. Roots and soils from grassland and forest in Hunan, China (27.944 °N, 113.226 °E) were sampled on 13 July 2021. Soil from the top 0-5 cm of the paddy field was collected using shovels. All samples were placed in sterile bags, kept on ice, and transported to the laboratory for further processing.

Three plant species (*Oryza sativa* indica, *Houttuynia cordata* and *Camphora officinarum*) each with triple replicates each were divided into three groups depending on the niche: root, rhizosphere, and bulk soil. The excess soil was vigorously shaken away from the roots to remove loose soil, leaving approximately 1 mm of soil also attached to the roots, which constitutes the rhizosphere compartment [1]. To separate the 1 mm of rhizosphere soil from the roots directly, roots were cut into 5 cm lengths and two grams were placed in a sterile flask with 50 mL of sterile phosphate buffered saline (PBS) solution and vortexed (15 seconds) several times. We avoided collecting any roots that were on the soil interface. The PBS solution used to wash the roots was poured into a 50 mL Falcon tube, centrifuged 10,000 *xg* for 1 minute, and then discarded the supernatant. The remaining soil samples were saved as rhizosphere samples [1]. The roots were then resuspended and vortexed three times (15 seconds each) before being transferred to new Falcon tubes as root samples. The soil far from the roots was stored as bulk soil.

Bacterial isolation

Root, rhizosphere, and bulk soil samples were put in the chamber for 3 to 4 days at 28°C

and 70% relative humidity on a 16:8-h day: night cycle. All root samples were washed with sterile milli-Q water to remove any adhering soil before they were subjected to the surface sterilization procedure described below: immersing samples in 75% ethanol for 3 minutes, then transferring to 5% sodium hypochlorite for 5 minutes, and finally rinsing samples 5 times with sterile distilled water [2]. The samples were then aseptically dried on the clean bench. The washed root samples were ground in a sterilized mortar and pestle with PBS buffer (pH 6.8). A fivefold and tenfold dilution was performed and then inoculated (100 µL) on modified arabinose-gluconate media without any nitrogen source [3]. To prepare soil inoculations, 5.0 g of fresh soil was mixed with 45 mL of sterile deionized water in a 50 mL Falcon tube. One mL of soil suspension was serially 10-fold diluted after mixing with a vortex mixer, and 100 µL diluted samples were also spread on the MAG media [3]. Then the isolation plates were placed in a 28 °C incubator to allow *Bradyrhizobium* to grow. After one week, small, white, or pink raised morphology formed were picked and repeatedly re-streaked on fresh MAG media for purification to obtain pure cultures. Purified strains were kept at -80 °C in glycerol suspensions (30% v/v).

The 10% (w/v) Chelex solution [Bio-rad, USA] and strain colonies were mixed and incubated in 60 °C for 20 minutes to extract DNA from bacterial colonies for Polymerase chain reaction (PCR) templates. The universal bacteria 16S rRNA primers 27F (5′-AGRGTTYGATYMTGGCTCAG-3′) and 1492R (5′-GGYTACCTTGTTACGACTT-3′), Premix Taq [Takara Bio, USA] and RNA/DNAse- free water were used in the PCR. The PCR conditions were as follows: denaturation at 95 °C for 5 minutes, followed by 35 cycles (95 °C for 45 s, 55 °C for 45 s and 72 °C for 90 s) and a final extension at 72 °C for 10 minutes [3]. PCR products were evaluated on 0.8% agarose and viewed by the automatic analysis system of electrophoresis gel

imaging. Positive PCR products were sequenced in BGI Genomics. The taxonomic information of each isolate was analyzed by comparing the 16S rRNA gene sequences with *Bradyrhizobium* in EzBioCloud [4].

Basic soil characteristics

Soil samples were separated to measure the basic soil characteristics after all the samples were returned to the laboratory. The soil water content, pH, soil organic carbon (SOC), total nitrogen (TN), total phosphorus (TP), dissolved organic carbon and nitrogen (DOC and DON), microbial biomass carbon and nitrogen (MBC and MBN), dissolved inorganic nutrients (i.e., nitrate, nitrite, ammonium, and phosphoric acid salt) and heavy metal content (Fe and Mo) were determined in the laboratory. These sample-associated metadata were provided in Dataset S1.

Genome sequencing, assembly and annotation

Bacteria Genomic DNA Extraction Kit [OMEGA Bacterial DNA Kit D3350] was used to extract genomic DNA from each of the 208 isolates. NanoDrop™ 2000 [Thermo Fisher, USA] was used to assess the quality of extracted DNA samples with the criteria A260/A280 > 1.8, A260/A230 > 2.0 and A260/A270 > 1.0. Whole genome sequencing was performed in Wuhan Huada Gene Biotechnology Company by using the MGISEQ-2000 PE150+150+10+10 platform. BBMerge in the BBmap package v38.79 was used to identify the untrimmed adapters associated with raw reads [5]. Trimmomatic v0.39 was then used to trim adapters and low-quality reads [6]. Reads with less than 40 bp in length were discarded, and the quality of the remaining reads was assessed by FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Contigs were assembled using SPAdes v3.10.1 based on the remaining high-quality paired-end reads with

default parameters [7]. Those contigs with lengths greater than 1,000 bp and a k-mer coverage over five were kept for further analysis. The quality of genome assemblies was assessed by CheckM v1.0.7 [8]. In general, 140 newly sequenced strains clustered with the previously sequenced strain STM3843 as Clade 1, and 68 clustered with the other 12 publicly available strains into Clade 2. Clade 1 members possessed a significantly lower GC content (63.67±0.07% vs. 65.39 ± 0.22%, $p < 0.001$, phylogenetic ANOVA) and similar genome size (7.45±0.14 Mb vs. 7.56±0.33 Mb, $p = 0.746$, phylogenetic ANOVA) compared with Clade 2 (Datasets S1 and S3).

We noted an assembly error involving the *nif* island that ought to harbor two *nifH* genes. This is because of the two copies of *nifH* genes with nearly identical sequences. It is known that genes with highly similar sequences are likely to cause such mis-assembly for short-read sequencing like Illumina or MGISEQ-2000 [9]. To check this issue, we mapped all raw reads to the assembled genome, and found that the sequencing depth of the *nifH* gene was twice that of other genes in the surrounding regions. We further performed Nanopore sequencing of three phylogenetically distantly related PB strains namely SZCCHNR3119, SZCCHNS2021 and SZCCHNS1050, and had their genomes assembled using Flye v2.6 [10]. Analysis of their genome sequence confirmed that these strains indeed have 2 *nifH* separated by roughly 43 genes. Therefore, we used the reference-guided method [11] to re-assemble the genomes of the remaining PB strains with the parameter of "--untrusted-contigs" in Spades v3.10.1 [7]. For each assembled genome, genes were predicted using Prokka v1.12 [12] and annotated using the Cluster of Orthologous Genes(COG) database [13] and Kyoto Encyclopedia of Genes and Genomes (KEGG) database [14, 15].

Phylogenomic tree construction for all available *Bradyrhizobium* lineages

To identify the phylogenetic distribution of our isolates in the genus *Bradyrhizobium*, a phylogenomic tree of *Bradyrhizobium* was built using IQ-Tree v2.2.0 [16] with our 208 isolates and 566 public genomes (outgroup included) downloaded from the NCBI Genbank database based on 123 shared single-copy genes identified in a previous phylogenomics study of *Bradyrhizobium* [3]. The parameter "-s alignment -spp partition -m MFP -mset LG,WAG,JTT -mrate E,G,I,G+I -bb 1000" was applied so that each gene was allowed to has its own best-fit substitution model automatically selected by ModelFinder implemented in IQ-Tree [16]. The branch support was assessed by 1000 ultrafast bootstrap approximations [17].

Phylogenomic and comparative genomic analyses for photosynthetic supergroup of *Bradyrhizobium*

OrthoFinder v2.3.4 [18] was used to identify orthologous gene families for all of the photosynthetic *Bradyrhizobium* strains. Each of the identified 3,205 single-copy ortholog families was aligned at the amino acid level using MAFFT v7.487 [19], and trimmed using trimAl v1.4.rev15 with parameters "-automated1 -resoverlap 0.55 - seqoverlap 60" [20]. A maximum likelihood phylogenomic tree of the PB supergroup was built based on the concatenated alignment of 3,205 single-copy orthologs using IQ-TREE v2.2.0 with 1,000 ultrafast bootstrap replicates [17]. The best-fit evolutionary model for each ortholog was determined by ModelFinder implemented in IQ-TREE [16].

As shown in Fig. S1, the Kakadu supergroup represents a sister lineage of the photosynthetic supergroup, but these two supergroups were connected by long branches, suggesting that the Kakadu supergroup was not an appropriate outgroup to root the species tree

of the photosynthetic supergroup. Therefore, the photosynthetic supergroup phylogenomic tree was rooted using the outgroup-independent rooting method, the minimum variance (MV) [21] and the minimal ancestor deviation (MAD) method [22]. MV seeks the root with the minimum variance of root-to-tip distances. MAD takes into consideration all branches as plausible root positions, calculates for each candidate the relative deviation from the clock-likeness, and determines the root with the minimal mean relative deviation from the molecular clock interpretation of all branches [22]. The constructed trees based on these two outgroup-independent methods indicated the same root position (Fig. 1 and Fig. S7A). To characterize the similarity between photosynthetic *Bradyrhizobium* at the genome content level, a genome content dendrogram was constructed using IQ-TREE (settings: -m MK -st MORPH) based on the presence and absence patterns of orthologous families identified using OrthoFinder (Fig. 1 and Fig. S7B). Phylogenetic ANOVA [23] was performed to compare continuous traits of strains while accounting for the evolution of those traits using the 'phylANOVA' function with 1,000 simulations implemented in the R package 'phytools' [24].

The similarity of the photosynthetic *Bradyrhizobium* strains was measured at both 16S rRNA gene and whole-genome levels. For the former, pairwise 16S rRNA gene identity was calculated using BLAST [25], and strains were clustered using the complete linkage method based on sequence identity. For the latter, the whole-genome average nucleotide identity was estimated by FastANI v1.2 [26].

Population genomics of photosynthetic *Bradyrhizobium*

PopCOGenT [27] was used to delineate genetically isolated populations for 221 *Bradyrhizobium* genomes (13 public and 208 new genomes) distributed in photosynthetic supergroup, to identify seven main clusters (MCs). The rationale behind this method is that as

bacterial speciation often develops rapidly, population boundaries should be delineated based on recent gene flow events [27]. One significant advantage of PopCOGenT is the ability to distinguish between recent and historical transfer events [27].

Symbiotic analysis on *Aeschynomene indica*

The 27 strains used for symbiotic analysis were grown in AG medium [28] at 28°C on Petri dishes for one week. Bacteria were harvested from the plate and resuspended in 10 mL of sterile water and the OD_{600} was adjusted to 1.0. *A. indica* plants were cultured as previously described [29]. Four plants per strain were inoculated with 1mL of bacterial suspension and symbiotic properties were analyzed at 17 days after inoculation (number of nodules per plant and nitrogenase enzyme activity) [30]. Cytological analysis of nodules elicited by one strain representative of the populations (MC2 to MC6) and three strains for the MC1 population was performed as described [31]. Each strain was tested in duplicate.

In vitro nitrogenase enzyme activity

Bacteria were grown in 9 mL vacuette® tubes (Greiner Bio-One GmbH) containing 2 mL of 0.8% agar BNM-B medium with 10 mM succinate and 10 mM arabinose as carbon source at 28 °C [32]. At the start of the experiment 10% acetylene was added to the vacuette® tubes and after eight days of incubation, the amount of ethylene produced by the bacterial culture was measured by gas chromatography [33].

Methanol metabolism

To compare the methanol utilization ability between PB populations, we selected 17

strains covering all populations predicted by PopCOGenT. The reference strain *Bradyrhizobium diazoefficiens* USDA110 was used as a positive control. All selected strains were inoculated into 14 mL test tubes containing 5 mL minimal salts medium, which were supplemented with 0.5% methanol as a sole carbon source and with 30 μM $CeCl_2$ (lanthanide chlorides) [34], and were incubated at 28 °C with reciprocal shaking at 280 rpm. The growth was monitored spectrophotometrically by measuring the optical density at 600 nm (OD_{600}).

Supplementary references

1. Edwards J, Johnson C, Santos-Medellin C, Lurie E, Podishetty NK, Bhatnagar S, et al. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci U S A* 2015; **112:** E911-920.
2. Mbai F, Magiri E, Matiru V, Nganga J, Nyambati V. Isolation and characterization of bacterial root endophytes with potential to enhance plant growth from Kenyan Basmati rice. *American International Journal of Contemporary Research* 2013.
3. Tao J, Wang S, Liao T, Luo H. Evolutionary origin and ecological implication of a unique nif island in free-living Bradyrhizobium lineages. *ISME J* 2021; **15:** 3195-3206.
4. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, et al. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017; **67:** 1613-1617.
5. Bushnell B, Rood J, Singer E. BBMerge–accurate paired shotgun read merging via overlap. *PLOS ONE* 2017; **12:** e0185056.
6. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30:** 2114-2120.

202 7. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a
203 new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*
204 2012; **19:** 455-477.
205 8. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the
206 quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome*
207 *Res* 2015; **25:** 1043-1055.
208 9. Tørresen OK, Star B, Mier P, Andrade-Navarro MA, Bateman A, Jarnot P, et al. Tandem
209 repeats lead to sequence assembly errors and impose multi-level challenges for genome and
210 protein databases. *Nucleic Acids Research* 2019; **47:** 10994-11006.
211 10. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using
212 repeat graphs. *Nat Biotechnol* 2019; **37:** 540-546.
213 11. Cabuk U, Unlu ES. A combined de novo assembly approach increases the quality of
214 prokaryotic draft genomes. *Folia Microbiol (Praha)* 2022; **67:** 801-810.
215 12. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30:** 2068-
216 2069.
217 13. Galperin MY, Wolf YI, Makarova KS, Vera Alvarez R, Landsman D, Koonin EV. COG
218 database update: focus on microbial diversity, model organisms, and widespread pathogens.
219 *Nucleic Acids Research* 2020; **49:** D274-D281.
220 14. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for
221 deciphering the genome. *Nucleic Acids Res* 2004; **32:** D277-280.
222 15. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for
223 taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research* 2022; **51:** D587-
224 D592.

16. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* 2020; **37:** 1530-1534.

17. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* 2017; **35:** 518-522.

18. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019; **20:** 238.

19. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013; **30:** 772-780.

20. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009; **25:** 1972-1973.

21. Mai U, Sayyari E, Mirarab S. Minimum variance rooting of phylogenetic trees and implications for species tree reconstruction. *PLOS ONE* 2017; **12:** e0182238.

22. Tria FDK, Landan G, Dagan T. Phylogenetic rooting using minimal ancestor deviation. *Nat Ecol Evol* 2017; **1:** 193.

23. Garland T, Dickerman AW, Janis CM, Jones JA. Phylogenetic Analysis of Covariance by Computer-Simulation. *Systematic Biology* 1993; **42:** 265-292.

24. Revell LJ. phytools 2.0: An updated R ecosystem for phylogenetic comparative methods (and other things)2023**:** 2023.2003.2008.531791.

25. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, et al. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* 2013; **41:** W29-33.

26. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018; **9:**

248 5114.
249 27. Arevalo P, VanInsberghe D, Elsherbini J, Gore J, Polz MF. A Reverse Ecology Approach
250 Based on a Biological Definition of Microbial Populations. *Cell* 2019; **178:** 820-834 e814.
251 28. Sadowsky MJ, Tully RE, Cregan PB, Keyser HH. Genetic Diversity in Bradyrhizobium
252 japonicum Serogroup 123 and Its Relation to Genotype-Specific Nodulation of Soybean. *Appl*
253 *Environ Microbiol* 1987; **53:** 2624-2630.
254 29. Okazaki S, Tittabutr P, Teulet A, Thouin J, Fardoux J, Chaintreuil C, et al. Rhizobium-
255 legume symbiosis in the absence of Nod factors: two possible scenarios with or without the
256 T3SS. *ISME J* 2016; **10:** 64-74.
257 30. Bonaldi K, Gourion B, Fardoux J, Hannibal L, Cartieaux F, Boursot M, et al. Large-scale
258 transposon mutagenesis of photosynthetic Bradyrhizobium sp. strain ORS278 reveals new
259 genetic loci putatively important for nod-independent symbiosis with Aeschynomene indica. *Mol*
260 *Plant Microbe Interact* 2010; **23:** 760-770.
261 31. Songwattana P, Chaintreuil C, Wongdee J, Teulet A, Mbaye M, Piromyou P, et al.
262 Identification of type III effectors modulating the symbiotic properties of Bradyrhizobium vignae
263 strain ORS3257 with various Vigna species. *Sci Rep* 2021; **11:** 4874.
264 32. Nouwen N, Arrighi J-F, Cartieaux F, Chaintreuil C, Gully D, Klopp C, et al. The role of
265 rhizobial (NifV) and plant (FEN1) homocitrate synthases in Aeschynomene/photosynthetic
266 Bradyrhizobium symbiosis. *Scientific Reports* 2017; **7:** 448.
267 33. Giraud E, Hannibal L, Fardoux J, Vermeglio A, Dreyfus B. Effect of Bradyrhizobium
268 photosynthesis on stem nodulation of Aeschynomene sensitiva. *Proc Natl Acad Sci U S A* 2000;
269 **97:** 14795-14800.
270 34. Wang L, Suganuma S, Hibino A, Mitsui R, Tani A, Matsumoto T, et al. Lanthanide-

dependent methanol dehydrogenase from the legume symbiotic nitrogen-fixing bacterium Bradyrhizobium diazoefficiens strain USDA110. *Enzyme Microb Technol* 2019; **130:** 109371.

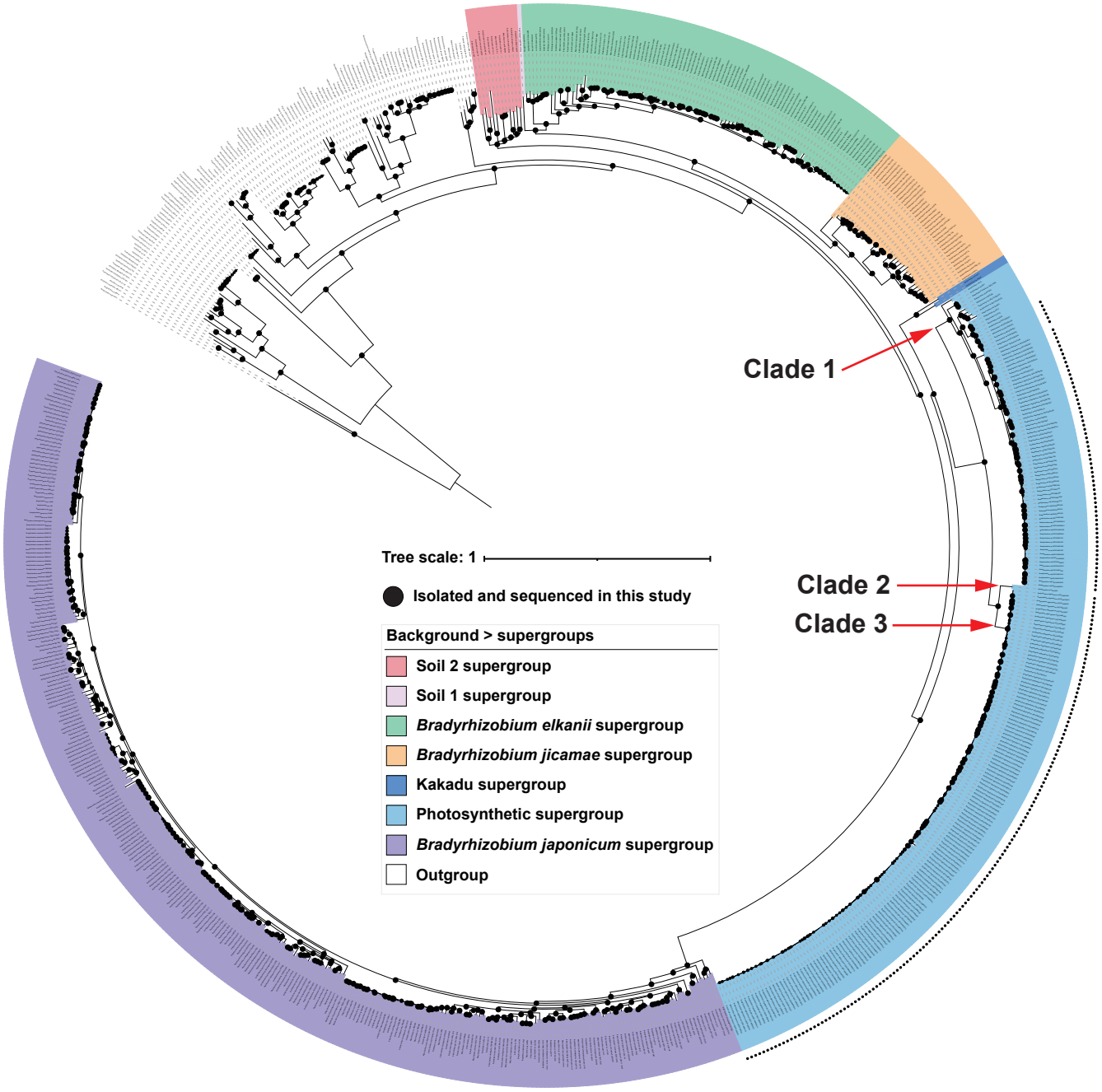


Figure S1. The maximum-likelihood phylogenomic tree of *Bradyrhizobium*. Strains from Xanthobacteraceae were used as an outgroup. The tree was constructed using the 123 orthologous genes identified in a previous study (Tao et al., 2021). The black circles on the nodes indicate ultrafast bootstrap values higher than or equal to 95% calculated by IQ-Tree. The 208 PB strains sequenced in the present study are indicated by black dots in the outermost layer surrounding the tree.

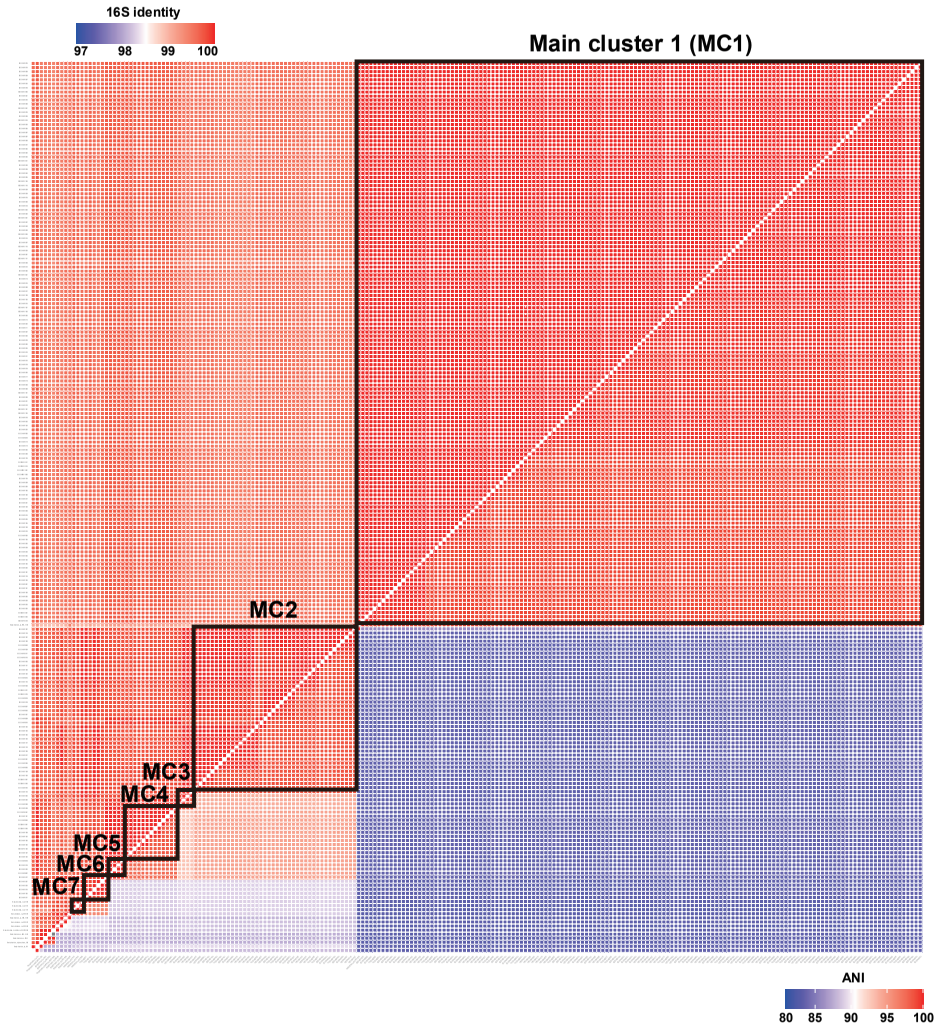


Figure S2. The heatmap of the pairwise identity of 16S rRNA genes and the whole-genome average nucleotide identity (ANI) of all PB members.

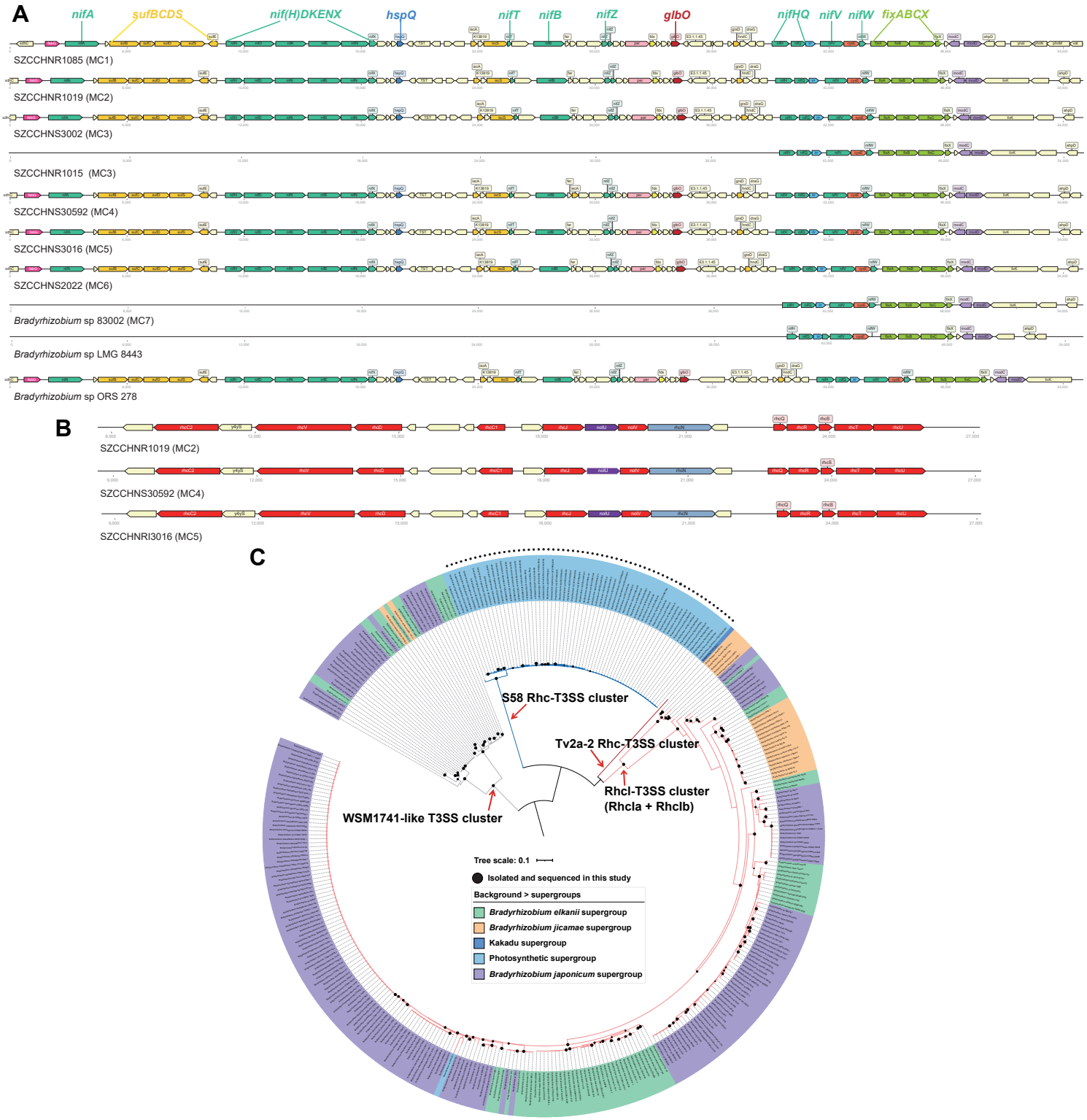


Figure S3. Comparison of the genomic context of the *nif* gene cluster (*nif* island) (A) and T3SS gene cluster (B) in the representative strains of PB of *Bradyrhizobium*. Gene functions are distinguished by different colors. The visualization of gene arrangement is performed with DNA-features-viewer v3.0.3 (Zulkower and Rosser, 2020). (C) The phylogenetic tree of the *rhcN* protein from *Bradyrhizobium*. The *rhcN* families were defined according to Teulet et al. (2020). The gene tree was rooted using the minimum variance (MV) method. The different colored branches correspond to the distinct genetic organization of the T3SS clusters to which the *rhcN* gene belongs. The *rhcN* in the strain *Bradyrhizobium* sp. 36 is not shown as it belongs to a different type of T3SS, likely a result of HGT from distantly related bacteria according to Teulet et al. (2020). Black circles in the phylogeny indicate nodes with IQ-Tree's ultrafast bootstrap values ≥ 95%.

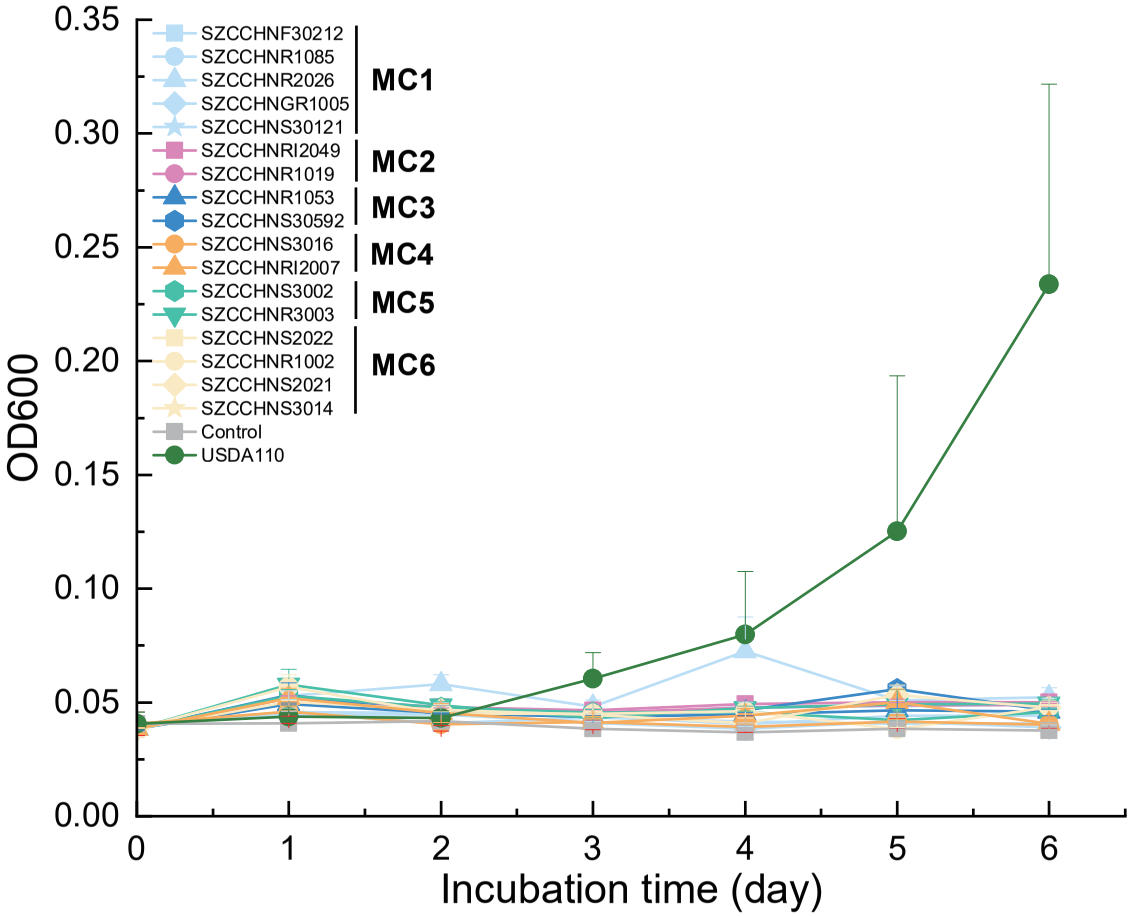


Figure S4. Growth of representative strains from the populations delineated by PopCOGenT for the PB of *Bradyrhizobium* on methanol as a sole carbon source, with the presence of lanthanide (Ln) species (Ce^{2+}, 30 uM). The reference strain *Bradyrhizobium diazoefficiens* USDA110 was used as a positive control. Error bars indicate the standard deviation of the mean from three replicates.

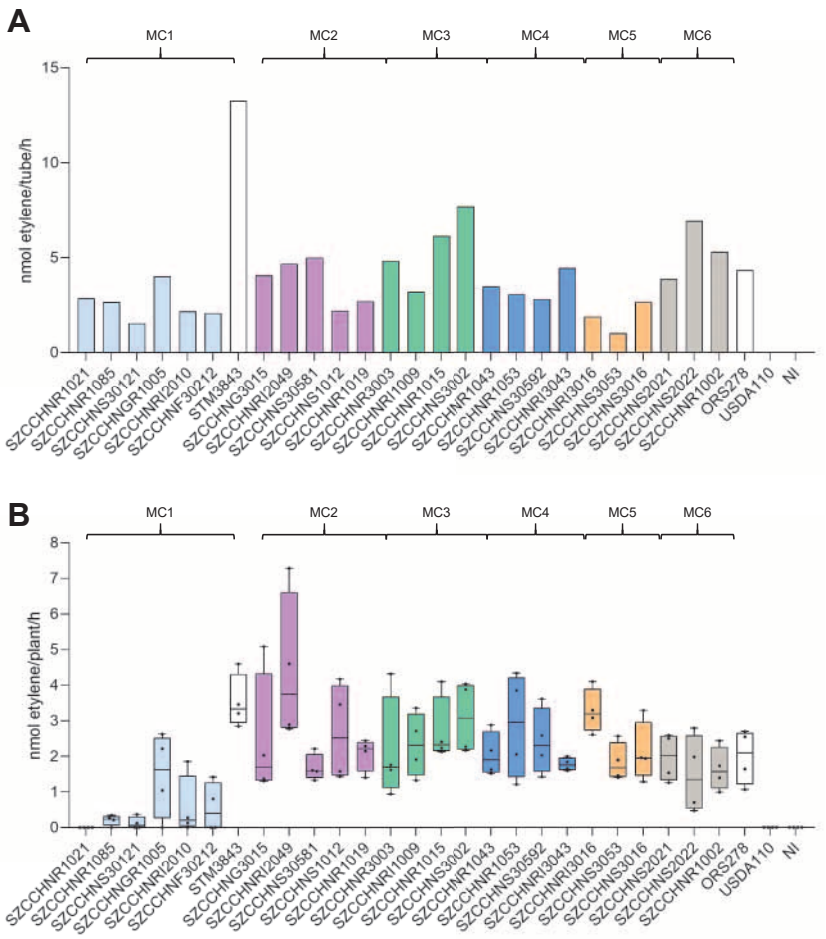


Figure S5. Free-living and symbiotic nitrogen fixation of several representative strains of the main populations (MCs) identified in the PB supergroup. (A) Free-living nitrogen fixation after 8 days of culture in vacuette® tubes. (B) Nitrogen fixation of *A. indica* plants inoculated with different PB representative strains at 17 days post-inoculation. In (A) and (B), ORS278 is used as a positive control and USDA110 (a member of *B. japonicum* supergroup) is used as a negative control. NI: non-inoculated.

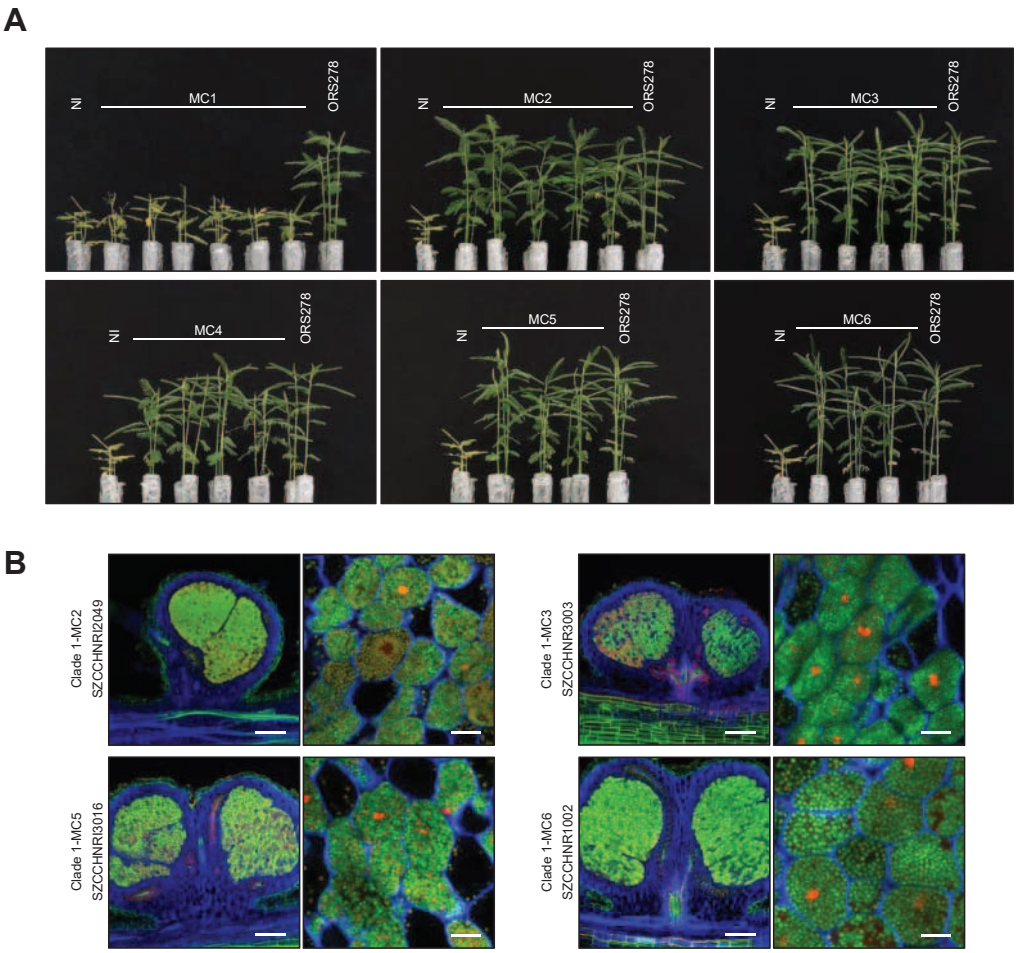


Figure S6. Complementary data of Fig. 2 showing symbiotic properties of several representative strains of the main populations identified in the PB supergroup. (A) Comparison of the growth of the *A. indica* plants (leaf phenotype) non-inoculated (NI) or inoculated with different representative strains of PB. All representative strains tested from each MC are present in this order: MC1 - SZCCHNR1021; SZCCHNR1085; SZCCHNS30121; SZCCHNGR1005; SZCCHNRI2010; SZCCHNF30212; MC2 - SZCCHNG3015; SZCCHNRI2049; SZCCHNS30581; SZCCHNS1012; SZCCHNR1019; MC3 - SZCCHNR3003; SZCCHNRI1009; SZCCHNR1015; SZCCHNS3002; MC4 - SZCCHNR1043; SZCCHNR1053; SZCCHNS30592; SZCCHNRI3043; MC5 - SZCCHNRI3016; SZCCHNS3053; SZCCHNS3016 and MC6 - SZCCHNS2021; SZCCHNS2022; SZCCHNR1002. ORS278 is used as control. (B) Confocal microscopy images of micro-sections of nodules elicited by the other Clade 1 strains tested after staining with SYTO9 (green, live bacteria), propidium iodide (red, infected plant nuclei and dead bacteria or bacteria with compromised membranes) and calcofluor (blue, plant cell wall). Scale bars: column 1, 200 µm; column 2, 10 µm.

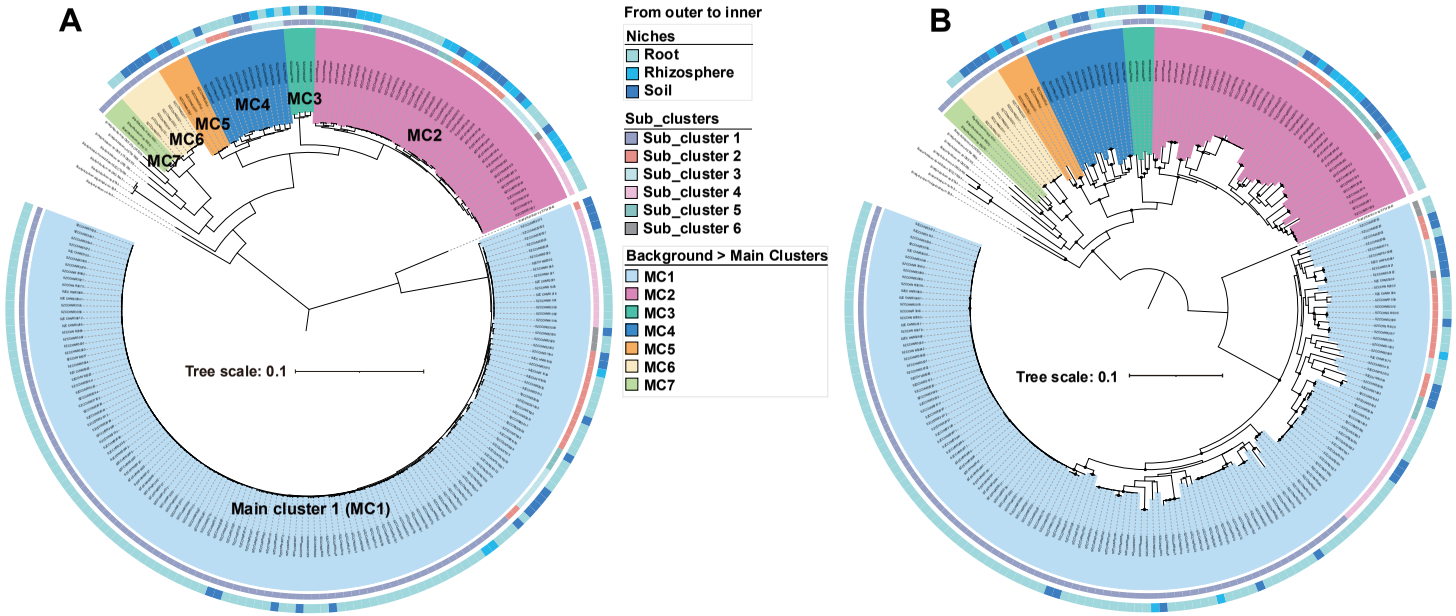


Figure S7. (A) The phylogenomic tree of the photosynthetic *Bradyrhizobium* is based on the minimal ancestor deviation (MAD) rooting method. (B) The gene content dendrogram of the photosynthetic supergroup of *Bradyrhizobium*. Solid circles in the phylogeny indicate nodes with IQ-Tree's ultrafast bootstrap values ≥ 95%.

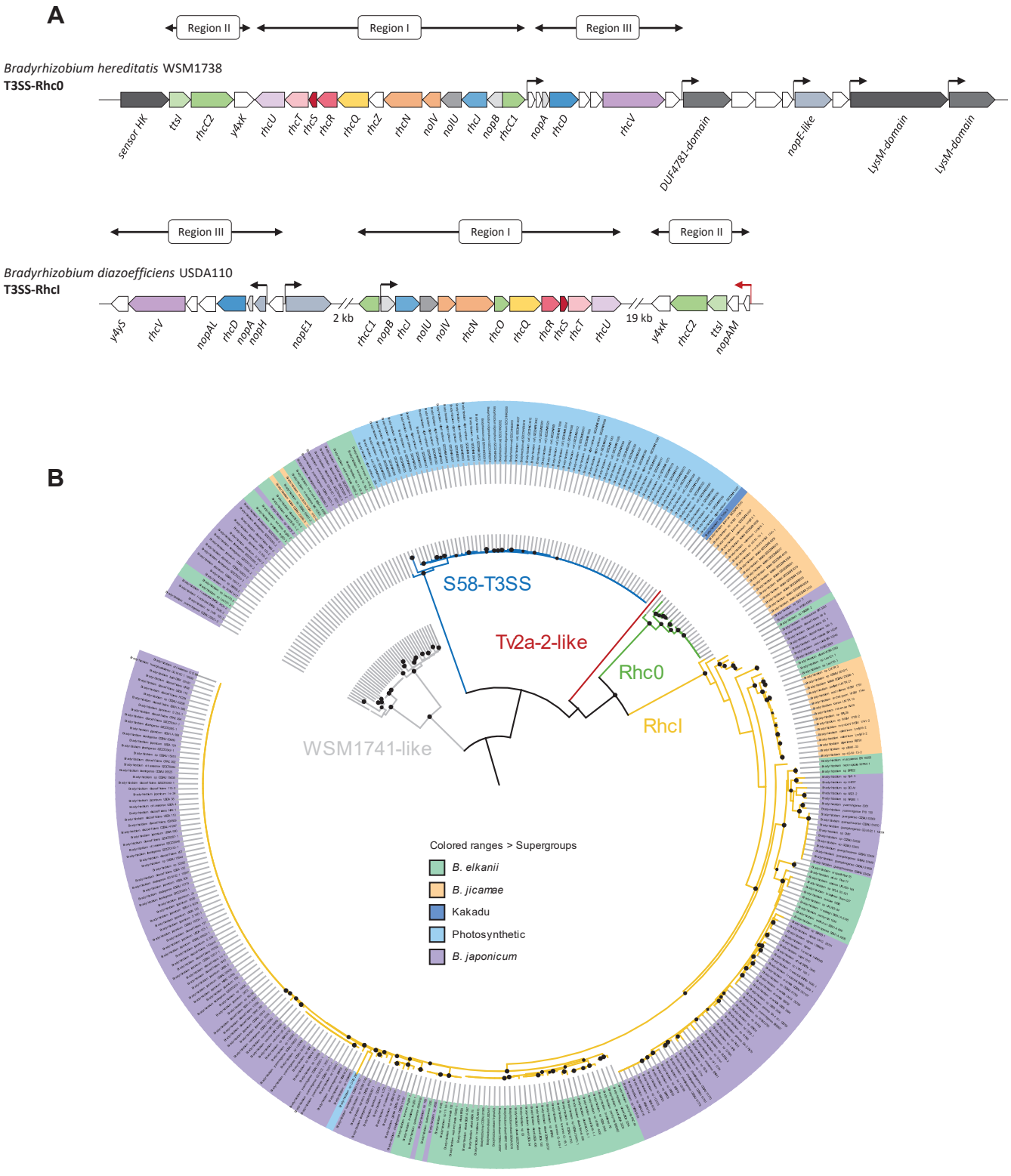


Figure 72. Rhc0 : un nouveau type de T3SS identifié chez les *Bradyrhizobium*.
(**A**) Organisation génétique du cluster Rhc0 et RhcI avec trois régions généralement conservées (I à III). Les gènes sont représentés par des flèches orientées selon le sens de la transcription et colorées selon les couleurs des différents composants du T3SS définis dans Teulet et al., 2021. Les gènes en gris foncé sont ceux apparaissant spécifiques au Rhc0 (sensor HK pour histidine kinase et certains T3Es putatifs identifiés par HMM et contenant un domaine DUF4781 ou LysM. Les gènes en blanc sont soit des T3Es déjà connus (ex: nopAL) ou des gènes de fonction inconnue (ex: y4xK). Les flèches plus fines indiquent les promoteurs avec une *tts* box (noir) ou une *nod* box (rouge). (**B**) Arbre phylogénétique de la protéine RhcN au sein du genre Bradyrhizobium en accord avec la nomenclature établie par Teulet et al., 2020. L'arbre a été enraciné en utilisant la méthode de la variance minimale (communication personnelle Lu Ling). La couleur des branches indique les différents cluster T3SS identifiables auxquels appartient le gène *rhcN*. Les cercles noirs indiquent les nœuds avec des valeurs de bootstrap ≥ 95 %.

3. Données complémentaires : de nouveaux T3SS chez les *Bradyrhizobium* dont le rôle reste à déterminer

La disponibilité de nouvelles souches de *Bradyrhizobium* séquencées récemment par le groupe de Haiwei Luo a permis d'identifier chez certaines bactéries un nouveau type de cluster de gènes T3SS qui est différent des six types précédemment identifiés (Teulet et al., 2020). A noter que jusqu'à présent seul le type RhcI, qui est le plus couramment retrouvé et quasi-systématiquement associé aux gènes *nod*, a été montré comme jouant un rôle symbiotique. De manière intéressante, ce nouveau type de T3SS est spécifiquement retrouvé au sein de souches appartenant au supergroupe *B. jicamae* qui ont été toutes isolées à partir de différents sols. Par ailleurs, l'ensemble des gènes présents dans le type RhcI incluant le régulateur transcriptionnel *ttsI*, est aussi retrouvé dans ce nouveau type de T3SS (figure 72A). Une analyse phylogénétique basée sur la protéine RhcN montre cependant que ce nouveau T3SS serait plus ancestral que le type RhcI (figure 72B) ; c'est pourquoi nous proposons de le nommer Rhc0. Un autre point remarquable est que les gènes *nod*, incluant *nodD* dont la protéine correspondante active normalement l'expression de *ttsI*, n'apparaissent pas associés à ce cluster Rhc0. En effet, la plupart des souches du phylogroupe *B. jicamae* contenant ce nouveau type de T3SS sont dépourvues de gènes *nod*, ce qui suggère un mode de régulation et une acquisition différente de ce cluster. Une analyse plus fine des gènes constituant le T3SS-Rhc0 conforte cette hypothèse d'un mécanisme de régulation différent. En effet, on observe la présence d'un gène codant pour un senseur histidine kinase retrouvé juste en amont de *ttsI* qui pourrait activer TtsI après transfert d'un groupement phosphate ; ce senseur étant absent chez les souches possédant le type RhcI. En accord avec cela, il peut être observé la présence d'un résidu d'aspartate au lieu d'un glutamate au niveau du site de phosphorylation chez les protéines TtsI identifiées dans les souches contenant un T3SS-Rhc0 (figure 73). Il avait été précédemment proposé par Marie et al., 2004, que le TtsI identifié dans les souches contenant un type RhcI ne nécessitait pas d'être phosphorylé pour être actif, suite justement à ce changement de résidu d'aspartate en glutamate, faisant de ce régulateur une forme constitutivement active.

Le régulateur TtsI étant conservé chez les souches possédant un T3SS-Rhc0, il peut également être possible de prédire les T3Es présents chez ces souches via la recherche d'un motif *tts* box. Une analyse réalisée chez l'une d'entre-elles (*B. hereditatis* WSM1738) permet de prédire quatre putatives *tts*-box en amont de gènes pouvant correspondre à des T3Es et qui sont retrouvés à proximité immédiate des gènes codant la machinerie du T3SS. Aucun de ces quatre gènes ne correspond à des T3Es de rhizobium connus et réciproquement, il est à noter qu'aucun homologue aux T3Es connus n'a été retrouvé chez ces souches par Blast. Parmi les gènes identifiés, deux codent pour des protéines contenant des domaines LysM qui sont connus pour permettre la reconnaissance de polysac-

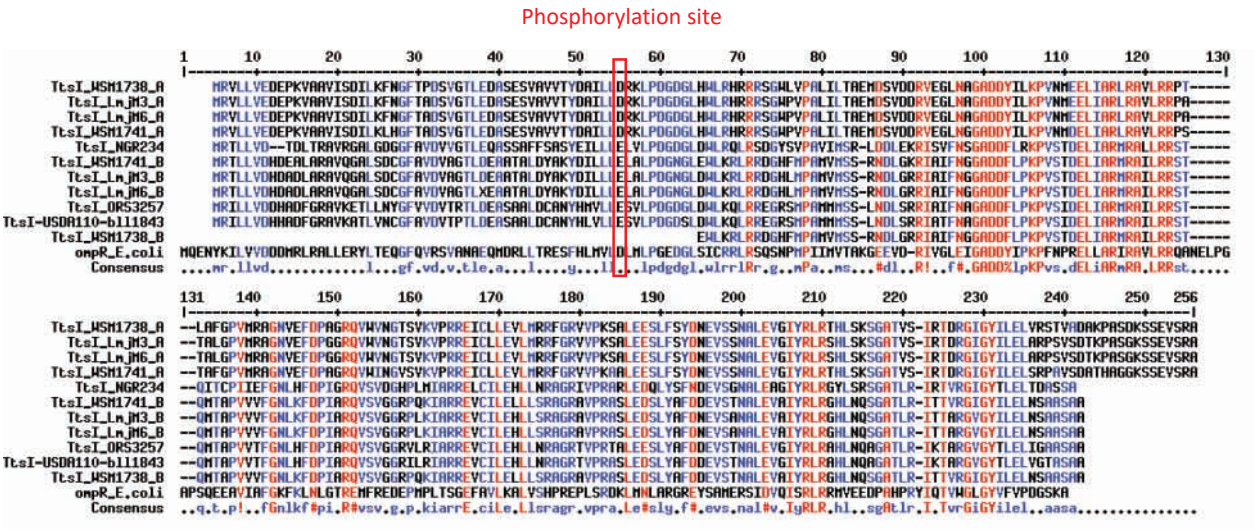


Figure 73. Le site de phosphorylation de TtsI varie selon le type de T3SS (Rhc0 ou RhcI).
Alignement de divers TtsI identifiés chez des souches de *Bradyrhizobium* contenant un Rhc0 (WSM1738, LmjM3, LmjM6, WSM1741) ou/et un RhcI (WSM1741, LmjM3, LmjM6, ORS3257 et USDA110), chez *Ensifer fredii* NGR234 et *E. coli*.

charides extracellulaires tels que le peptidoglycane ou la chitine. Ce type de T3Es à domaine LysM est surtout retrouvé chez les champignons et sont supposés les protéger d'une mise en place d'une réponse immunitaire de la plante induite par les chitines (Hu et al., 2021). En effet, la plante synthétise des chitinases en réponse au champignon qui dégradent leur paroi en libérant des composés de chitines pouvant être reconnu comme DAMPs par les récepteurs de la plante. On ne peut donc pas exclure que ces T3Es à domaine LysM jouent un rôle similaire mais en masquant les muropeptides[143] susceptibles d'être sécrétés par ces bactéries Gram négatives au cours de leur division cellulaire.

Un autre point remarquable qui ressort de l'analyse des génomes de ces nouvelles souches séquencées est la mise en évidence que le type T3SS-S58, initialement identifié chez une unique souche de *Bradyrhizobium* (*B. ologotrophicum* S58), est retrouvé chez près d'un tiers des nouvelles souches de *Bradyrhizobium* photosynthétiques isolées. La souche S58 qui appartient également à ce même clade, a été isolée d'un sol de rizière (Okubo et al., 2013), tout comme ces nouvelles souches séquencées et contenant un T3SS-S58. Il serait donc intéressant d'explorer si ce T3SS facilite l'interaction de ces souches avec le riz et s'il permet ainsi d'améliorer sa croissance ; d'autant plus qu'une étude récente conduite sur le T3SS d'une souche bénéfique du riz (*Burkholderia vietnamiensis* LMG10929) a montré que celui-ci joue un rôle majeur dans la colonisation du système racinaire du riz (Wallner, 2021). Le T3SS-S58, toujours basé sur l'étude phylogénétique du gène *rhcN*, apparaît encore plus ancestral que le Rhc0 (figure 72B). A noter également que les gènes correspondants à la piline NopA ou le translocon NopX ou NopE ne sont pas présents. Il se pose ainsi la question si ce T3SS permet véritablement la sécrétion de T3Es dans la plante hôte. On ne retrouve pas non plus d'homologue de *ttsI* et il n'est donc pas possible de prédire les T3Es susceptibles d'être sécrétés, au moins dans le milieu extérieur, par ces souches *via* la recherche de *tts* box. Par contre, on retrouve associé au cluster de gènes T3SS-S58 un gène codant une protéine hypothétique présentant également des domaines LysM et une faible similarité avec un T3E de type LysM putatif identifié chez les souches du phylogroupe *B. jicamae* contenant un Rhc0. On peut donc suggérer que ces T3Es à domaine LysM constituent les premiers T3Es originaux acquis chez les bactéries du genre *Bradyrhizobium*. Le fait que ce type de T3Es ne soit pas retrouvé chez les *Bradyrhizobium* possédant un T3SS de type RhcI est sûrement à relier avec le fait que ce cluster est systématiquement associé aux gènes *nod*. Il se peut donc que les T3Es de type LysM aient été perdus au cours de l'évolution pour ne pas interférer avec les FNs qui sont susceptibles d'être reconnus par ces T3Es à domaine LysM.

L'étude fonctionnelle de ces différents T3SS (Rhc0 et S58) plus primitifs que le type RhcI apparaît ainsi particulièrement intéressante à conduire dans le futur pour mieux comprendre comment ce

[143] Dérivés de peptidoglycanes.

système de sécrétion a pu évoluer au sein des rhizobiums ; et ainsi sûrement permettre en s'associant avec les gènes *nod* le passage de la vie libre à symbiotique.

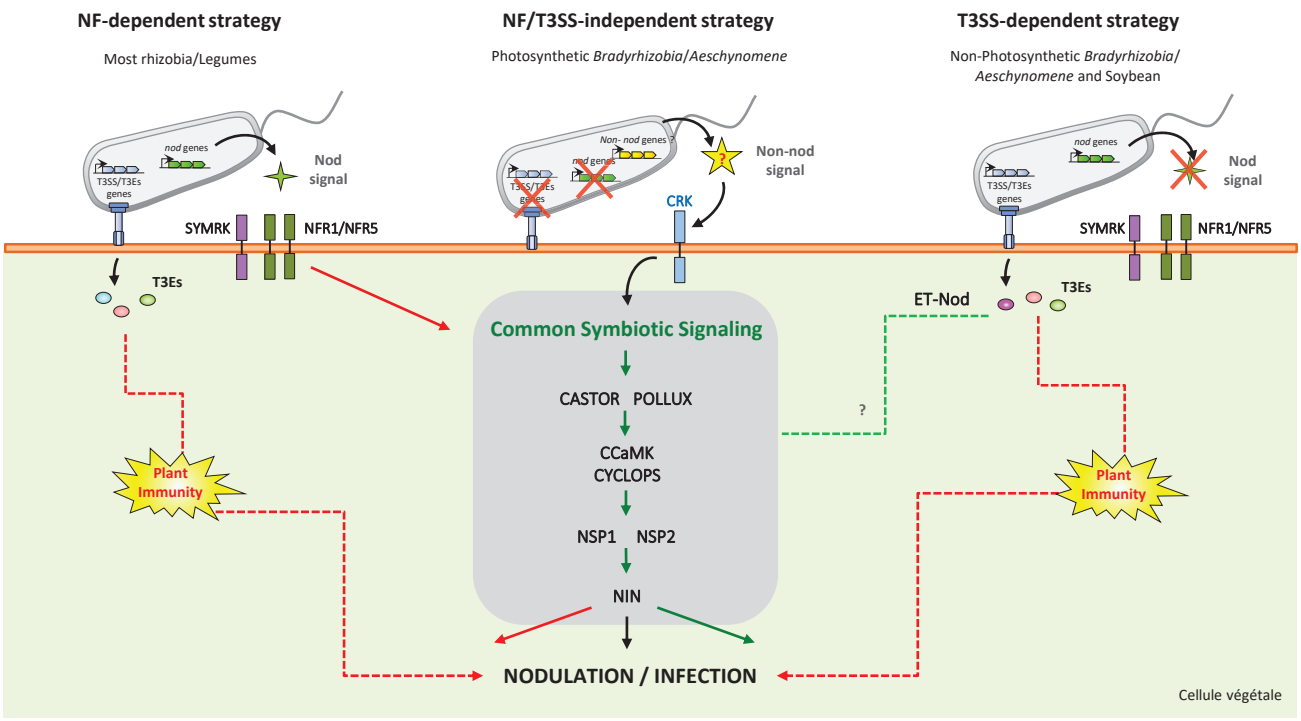


Figure 74. Diversité des processus symbiotiques entre rhizobiums et légumineuses.
Trois stratégies symbiotiques ont été identifiées à ce jour : un processus dépendant des FNs qui est le plus répandu (à gauche) et dont l'efficience peut être modulé ou pas par des T3Es injectés *via* le T3SS ; un autre indépendant à la fois des FNs et du T3SS qui est restreint à quelques espèces du genre *Aeschynomene* et des *Bradyrhizobium* photosynthétiques (au milieu) et un dernier, lui aussi indépendant des FNs mais où le T3SS joue un rôle central dans le cas d'interaction entre des *Bradyrhizobium* non-photosynthétiques et des espèces du genre *Aeschynomene* et le soja (à droite). Dans ce dernier cas, le T3SS permet l'injection d'un cocktail de T3Es ayant des rôles synergiques et complémentaires, dont les ET-Nods activant l'organogénèse nodulaire. Chacune de ces stratégies possède des acteurs communs (common symbiotic signalling) et des signaux associés à des récepteurs ou cibles spécifiques qui vont conduire ensemble à l'infection et l'organogénèse nodulaire.

Discussion et perspectives

L'ensemble des études conduites sur la symbiose rhizobium/légumineuse permet aujourd'hui de définir trois processus distincts qui mènent tous au développement d'un nouvel organe racinaire hébergeant les bactéries, le nodule (figure 74). Le premier mécanisme, qui est le plus répandu, repose sur une stratégie de nodulation et d'infection dépendante des FNs. Il est retrouvé chez la plus grande majorité des couples symbiotiques et dans ce cas-là, le T3SS retrouvé chez certains rhizobiums module positivement ou négativement l'interaction à travers la sécrétion de T3Es qui activent ou répriment le système immunitaire de la plante hôte. Il a par la suite été découvert deux autres processus symbiotiques qui ne dépendent pas de la sécrétion et de la reconnaissance de FNs. Le premier, retrouvé uniquement chez des *Bradyrhizobium* appartenant au phylogroupe dit "photosynthétique" et interagissant spécifiquement avec certaines espèces d'*Aeschynomene*, implique probablement un signal bactérien qualifié de "non Nod" reconnu par la plante à travers de nouveaux récepteurs symbiotiques qui restent à caractériser. Le second est quant à lui retrouvé chez une large diversité de *Bradyrhizobium* non-photosynthétiques interagissant avec le soja ou certaines espèces d'*Aeschynomene*. A l'inverse du premier processus Nod-indépendant, ce second processus est dépendant de la sécrétion *via* le T3SS d'effecteurs spécifiques, appelés ET-Nods. Ces deux processus symbiotiques alternatifs et originaux sont cependant peu étudiés en comparaison du processus Nod-dépendant. Il existe donc un réel intérêt à améliorer nos connaissances sur ces nouveaux processus symbiotiques. Nous pourrons ainsi avoir une meilleure compréhension de l'évolution des mécanismes régissant la symbiose rhizobium/légumineuse et aussi mieux optimiser cette interaction afin de favoriser le développement d'une agriculture plus respectueuse de l'environnement.

1. La symbiose Nod-indépendante, T3SS-indépendante : un processus très efficace mais restreint au supergroupe des *Bradyrhizobium* "photosynthétiques"

1.1. Nouvelles connaissances sur le phylogroupe des *Bradyrhizobium* "photosynthétiques"

Le processus symbiotique alternatif Nod-indépendant, T3SS-indépendant n'a été décrit à ce jour que chez certaines espèces d'*Aeschynomene* nodulées par des *Bradyrhizobium* photosynthétiques. Ce système symbiotique est parfaitement fonctionnel puisqu'il permet d'éliciter des nodules fixateurs d'azote dès 4 jours après inoculation, permettant d'améliorer fortement la croissance de la plante hôte en conditions limitantes d'azote disponible. Ces souches nodulant naturellement ces *Aeschynomene* spp. appartiennent toutes au même supergroupe phylogénétique, initialement nommé "photosynthétique" ou PB pour "Photosynthétiques *Bradyrhizobium*" (Avontuur et al., 2019). Aujourd'hui, le séquençage de plus de 200 nouvelles souches appartenant à ce phylogroupe, isolées

pour la plupart du riz ou de la rhizosphère du riz, montre que plusieurs sous-populations à l'intérieur de ce phylogroupe n'ont pas les gènes photosynthétiques. Il apparaît ainsi que ce qualificatif de "photosynthétique" pour nommer ce phylogroupe n'est plus approprié puisqu'il ne constitue pas un trait commun à l'ensemble des souches qui en font partie. D'autant plus que le cluster de gènes photosynthétiques a été retrouvé chez diverses souches de *Bradyrhizobium* (25) appartenant aux supergroupes *B. japonicum* et *B. elkanii* (Avontuur et al., 2023). En revanche, nous montrons que l'ensemble des représentants testés appartenant aux différentes sous-populations identifiées à l'intérieur de ce phylogroupe présente la capacité de noduler *A. indica*, bien que ces souches n'aient pas été isolées de nodule. Fait intéressant, on observe une absence des gènes *nod* chez l'ensemble de ces souches, à l'exception de *Bradyrhizobium* sp. ORS285 déjà décrite auparavant. Nous proposons ainsi de renommer ce phylogroupe en "*nod* free but nodulating *Bradyrhizobium*" ou "*Bradyrhizobium* nodulant en l'absence de gènes *nod*".

Il est à noter que depuis les travaux présentés dans l'article 3, d'autres souches de *Bradyrhizobium* isolées d'autres environnements (sols de forêt, sols cultivés pour d'autres céréales que le riz) ont été séquencées par le groupe d'Haiwei Luo. Plusieurs nouvelles souches ont été déterminées comme appartenant au supergroupe des PB. Parmi celles-ci, il a été identifié une deuxième souche du clade 2 qui n'était représenté jusqu'alors que par la souche *Bradyrhizobium* sp. STM3843, ainsi qu'un autre clade qui apparaît phylogénétiquement basal au groupe des BP. Des tests symbiotiques sont en cours pour évaluer leur capacité à noduler *A. indica* et ils devraient permettre de mieux comprendre l'évolution de cette capacité de nodulation au sein de ce phylogroupe.

1.2. Nouvelles pistes pour l'identification du signal "non Nod"

Bien que des différences soient constatées dans l'efficience symbiotique, les observations faites dans l'article 3 ont mis en évidence la capacité universelle des souches de *Bradyrhizobium* appartenant au phylogroupe des PB à noduler *A. indica* de manière Nod-indépendante. Ainsi, les gènes symbiotiques gouvernant ce processus symbiotique Nod-indépendant auraient pu être acquis chez un ancêtre commun à ce phylogroupe puis transmis de manière verticale à l'ensemble des descendants, qui auraient conservés ce caractère. Ce scénario évolutif est complétement différent de celui décrit pour les gènes *nod* qui ont été acquis à de multiples reprises de manière horizontale au cours de l'évolution et même parfois perdus. La conservation des gènes symbiotiques chez ces souches, alors que celles-ci ont été isolées de divers environnements sans aucune pression de sélection exercée par des *Aeschynomene*, suggère que ces déterminants pourraient jouer un rôle clé dans la survie de ces bactéries. Ceci expliquerait pourquoi les tentatives de mutagénèse "perte de fonction" développées à ce jour pour identifier les déterminants "non Nod" ont échoué.

L'identification du signal "non Nod" est l'un des objectifs majeurs de l'équipe MSLT depuis maintenant 15 ans. La diversité émergente du supergroupe des PB et le concept selon lequel ces gènes symbiotiques seraient essentiels à la survie de la bactérie nous donnent aujourd'hui de nouvelles pistes exploratoires pour rechercher ces signaux symbiotiques.

Dans un premier temps, des approches gain de fonction pourraient être développées. La construction d'une banque de cosmides contenant des fragments d'environ 30-40 kb de l'ADN génomique d'une souche nodulante, telle que *Bradyrhizobium* sp. ORS278, ou une autre souche aussi efficiente sur *Aeschynomene*, serait une stratégie envisageable. Ces divers fragments d'ADN, une fois transférés et insérés dans une souche réceptrice incapable de noduler, pourraient conférer la capacité de nodulation à certains clones si les déterminants gouvernant la symbiose Nod-indépendante, T3SS-indépendante sont colocalisés. Nous pourrions ainsi identifier les régions génomiques gouvernant ce processus symbiotique. Il conviendrait de tester différentes souches réceptrices, sachant que des facteurs d'incompatibilité déclenchant une réponse immunitaire forte chez la plante ne sont pas à exclure. Une souche basale au supergroupe des PB initialement incapable de noduler semblerait idéale, d'où l'intérêt des tests symbiotiques en cours. Pour favoriser la sélection de clones capables d'initier la nodulation, il pourrait être aussi envisagé de travailler sur un mutant d'*A. evenia* hypernodulant (tel que les mutants *sun* disponibles). Parallèlement, les souches du clade 3 qui ont été montrées moins efficaces pour la nodulation chez *A. indica*, pourraient aussi être utilisées comme souches réceptrices pour mieux comprendre pourquoi elles présentent ce phénotype altéré.

Une autre stratégie serait de cibler spécifiquement certains déterminants génétiques identifiables par analyse comparative du core-génome des BP avec les autres core-génomes des *Bradyrhizobium* appartenant aux autres phylogroupes. Cependant, d'autres approches que celle d'absence/présence de gènes seront sûrement à considérer pour permettre l'identification de ces déterminants. Le grand volume de données génomiques maintenant disponible, associé à l'utilisation de méthodes de Machine Learning ou Deep Learning, qui utilisent l'intelligence artificielle et qui pourraient intégrer d'autres critères tels que la vitesse évolutive des gènes, la synténie ou la composition en acides-aminés des protéines correspondantes, pourraient faire émerger des candidats insoupçonnés.

Dans un second temps, les régions génomiques et/ou gènes identifiés par les approches gain de fonction devront être validés comme requis pour la nodulation Nod-indépendante, T3SS-indépendante. Pour cela, leur mutagénèse dans la souche initiale nodulante permettra de confirmer leur rôle symbiotique si une perte de la nodulation est observée et si la physiologie de la bactérie change. Il ne serait pas surprenant que la mutation soit létale pour la bactérie si on considère notre hypothèse selon laquelle ces déterminants symbiotiques seraient nécessaires à sa survie. Une possibilité de redondance fonctionnelle n'est également pas à exclure, des doubles ou triples mutants seront peut-être nécessaires pour valider nos candidats. La nature des gènes qui seront identifiés nous guidera

ainsi sur les approches fonctionnelles appropriées à développer pour mieux comprendre la nature du signal "non Nod" déclenchant la nodulation. Enfin, les récentes connaissances acquises du côté plante avec l'identification de plusieurs récepteurs symbiotiques (AeCRK, AeRLCK, AeRINRK et AeRLP), qui feraient parties d'un complexe signalétique, pourraient nous aider à identifier des signaux de cette symbiose atypique en développant des approches biochimiques de type ligand-fishing[144]. Cependant, il est nécessaire pour cela de mieux caractériser ce complexe signalétique et, dans ce sens, des approches de marquage de proximité (Turbo-ID) ou de phosphorylome sont en cours.

2. La symbiose Nod-indépendante, T3SS-dépendante : un processus plus largement répandu chez les *Bradyrhizobium* mais dont le mutualisme reste à démontrer

Contrairement à la symbiose Nod-indépendante, T3SS-indépendante qui n'a été mise en évidence qu'au sein des BP interagissant avec les *Aeschynomene* spp., la capacité à développer une symbiose Nod-indépendante, T3SS-dépendante est quant à elle retrouvée chez une large diversité de *Bradyrhizobium* (article 1). Néanmoins, aucune souche testée jusqu'à présent et utilisant un processus T3SS-dépendant ne conduit à une symbiose bénéfique pour la plante, que ce soit avec *Aeschynomene* spp. ou avec le soja. L'absence d'une symbiose efficiente chez ces souches s'explique phénotypiquement par un défaut d'infection intracellulaire malgré l'induction de nodules, ou par une non-différenciation en bactéroïde (la forme active fixatrice d'azote) lorsque le symbiote est internalisé dans le nodule.
Ce processus symbiotique alternatif n'est donc pas si mutualiste qu'il n'y paraît. Selon la définition originale d'une symbiose énoncée dans la partie 1.2.2 de l'introduction, le processus Nod-indépendant, T3SS-dépendant répond bien au critère d'une "association vivante d'espèces différentes". Cependant, à la différence des deux autres processus symbiotiques mutualistes décrits dans la symbiose rhizobium/légumineuse, l'avantage que tirent les deux symbiotes durant cette interaction T3SS-dépendante reste à démontrer. Le niveau de fixation de l'azote par les bactéries hébergées dans le nodule, lorsqu'il est détectable, reste très faible et n'apparaît pas suffisant pour observer un gain de croissance de la plante par rapport à une plante non inoculée. Cette interaction semble donc se rapprocher plus du commensalisme puisqu'elle ne semble bénéficier qu'à un seul des deux symbiotes sans altérer l'autre. En effet, la bactérie possède une niche écologique au sein du nodule qui peut être favorable à son développement, à l'inverse de la plante qui ne tire pas avantage de cette association. L'absence de différenciation en bactéroïdes, normalement médiée par la production de NCR chez *Aeschynomene*, suggère que les bactéries internalisées dans le nodule peuvent à tout mo-

[144] Ou "pêche au ligand" ; technique qui repose sur la sélection par affinité d'un ligand à partir d'un échantillon biologique par un récepteur immobilisé.

ment retourner à la vie libre. Ces observations soulèvent les interrogations suivantes : que manque-t-il à ce processus symbiotique pour qu'il soit pleinement fonctionnel et, dans la nature, une symbiose dépendante uniquement du T3SS existe-t-elle vraiment ?

Dans le cas de l'interaction entre *B. elkanii* USDA61 et le soja, l'inefficience des nodules élicités par le processus T3SS-dépendant sur le mutant *nfr1* est d'autant plus troublante que la souche induit des nodules parfaitement fixateurs sur la même lignée sauvage (Okazaki et al., 2013). On peut se demander si les FNs interviennent dans ce cas-là dans les étapes plus tardives du processus symbiotique, ou si le mode d'infection qui apparaît alors différent (intracellulaire pour le processus Nod-dépendant, intercellulaire pour le processus T3SS-dépendant) implique des contrôles de reconnaissance de la bactérie par la plante plus strictes pour activer les étapes suivantes conduisant à la fixation biologique de l'azote par la bactérie. Ce défaut de fixation d'azote des nodules formés de manière T3SS-dépendante rappelle celui observé lors d'une évolution expérimentale conduite chez la légumineuse *Mimosa pudica* inoculée avec la bactérie pathogène *Ralstonia solanacearum*, dans laquelle le plasmide symbiotique de *Cupriavidus taiwanensis* a été transféré (Clerissi et al., 2018; de Moura et al., 2020). En effet, si après plusieurs dizaines de cycles successifs d'inoculation sur *M. pudica* et réisolation des bactéries des nodules, il a pu être obtenu plusieurs lignées bactériennes capables d'infecter intracellulairement les nodules, aucune d'entre-elles n'a pu évoluer jusqu'à un stade lui permettant de franchir cette dernière étape de fixation d'azote. Ce défaut de fixation, comme observé dans notre étude, pourrait dès lors avoir une origine commune qui reste à ce stade inconnue.

La question de l'existence d'une symbiose fonctionnelle dépendante du T3SS dans la nature demeure, étant donné que la plupart des souches de *Bradyrhizobium* séquencées et pour lesquelles un T3SS a pu être identifié présentent également les gènes *nod*. La présence quasi-systématique de ces gènes suggère que les FNs doivent jouer un rôle important durant l'interaction de ces bactéries avec leur plante hôte d'origine. Il existe cependant trois souches non-photosynthétiques qui constituent des exceptions, puisqu'elles possèdent uniquement un T3SS : *Bradyrhizobium* sp. Y36, *B. liaoningense* CCBAU83689 et *B. mercantei* SEMIA6399 (Teulet et al., 2020). Ces souches, bien que non isolées d'*Aeschynomene* spp., ne possèdent pas les gènes *nod* et ont pourtant été isolées de nodules. Dans le cas de SEMIA6399, sa capacité à noduler sa plante hôte (*Deguelia costata*[145]) semble indiquer que cette souche est capable de noduler de manière Nod-indépendante (Helene et al., 2017). Si l'identité de la souche séquencée et sa capacité à noduler sont bien confirmées, cette souche serait la première n'appartenant pas au supergroupe des PB à être capable de noduler sans FNs d'autres légumineuses qu'*Aeschynomene* spp. Il serait intéressant d'élucider le rôle symbiotique de son T3SS, qui présente un cluster atypique différent du RhcI largement retrouvé chez les *Bradyrhizobium*, et

[145] Synonyme : *Lonchocarpus costatus*.

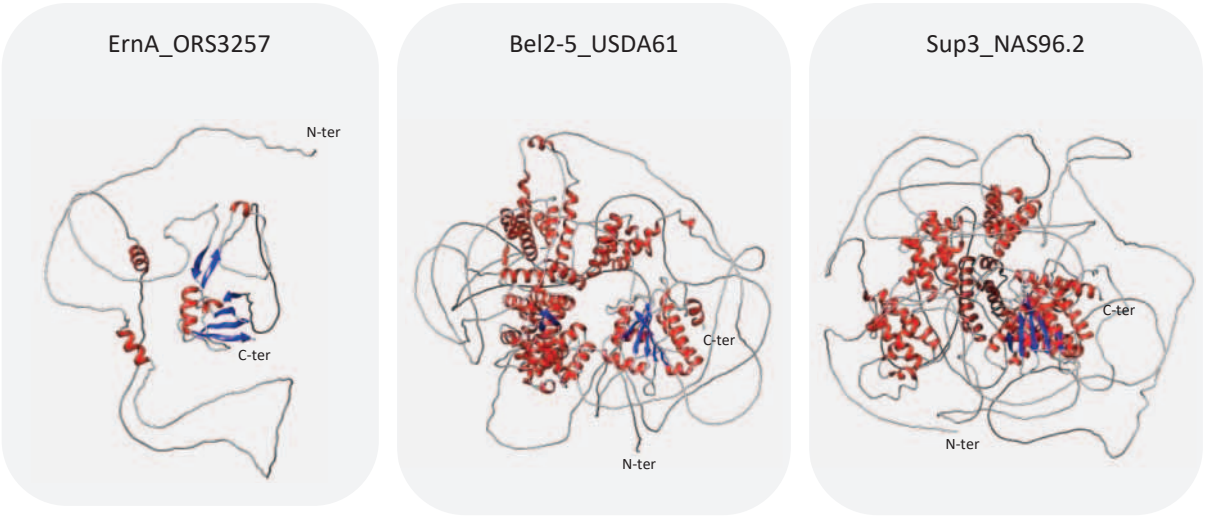


Figure 75. Prédiction structurale d'ErnA, Bel2-5 et Sup3 sur Alphafold.
Prédiction de la structure des trois ET-Nods qui ont été validés fonctionnellement (ErnA, Bel2-5 et Sup3) à partir de leur séquence protéique. Les structures secondaires apparaissent en rouge (hélices α) et en bleu (feuillet β). Le reste de la séquence pour laquelle aucune structure n'a pu être attribuée apparaît en gris et correspondrait aux séquences "en désordre" ou "disorder region" caractéristiques des effecteurs.

d'examiner également la capacité des BP à noduler cette légumineuse. Concernant les deux autres souches Y36 et CCBAU83689, il reste à confirmer si elles sont les vrais symbiotes de leur plante respective (*Glycine max* et *Phaseolus vulgaris*) et non des bactéries opportunistes, et si elles peuvent noduler de manière Nod-indépendante, T3SS-dépendante. L'ensemble de ces informations ne nous permet donc pas d'exclure qu'une symbiose rhizobium-légumineuse reposant uniquement sur le T3SS existe bien dans la nature.

3. Diversité des ET-Nods

3.1. Des effecteurs distincts les uns des autres mais qui partagent certaines similitudes

L'étude de la symbiose Nod-indépendant, T3SS-dépendante chez *A. indica* et le soja a permis de démontrer l'existence d'une nouvelle famille de T3Es capables de déclencher le programme de nodulation : les ET-Nods. Cette famille est constituée de trois effecteurs que sont ErnA, Bel2-5 et Sup3 ; plus trois putatifs ET-Nods (Ubi1, Ubi2 et Ubi3) dont la sécrétion et leur rôle fonctionnel n'ont pas encore été validés.

Lorsqu'on examine la séquence protéique de ces ET-Nods, il apparaît peu de similarités, bien que des points communs émergent. En effet, tous ces T3Es possèdent au moins un signal d'adressage au noyau (NLS) prédit dans leur séquence et des signatures communes ont pu être identifiées grâce au logiciel MEME, dont notamment le motif 2 qui présente une conservation des acides-aminés GERYTAxL retrouvés chez l'ensemble des ET-Nods identifiés (article 1). La localisation nucléaire d'ErnA, Bel2-5 et Sup3 a pu être démontrée chez le tabac ainsi que chez Ubi1 (communication personnelle Mélanie Carcagno), suggérant que cette famille de T3Es jouerait sur l'expression de gènes cibles pour déclencher la nodulation. L'importance du motif 2 pour l'activité fonctionnelle d'ErnA a par ailleurs été montrée dans l'article 1, ce qui laisse supposer qu'il pourrait jouer également un rôle important chez les autres ET-Nods.

La prédiction structurale des ET-Nods à l'aide d'Alphafold (Mirdita et al., 2022) n'a pas non plus permis d'identifier des similarités de structures, la prédiction de la structure des effecteurs étant peu concluante (figure 75). Cela peut s'expliquer par la nécessité des T3Es de posséder une structure flexible[146] afin de pouvoir se dérouler rapidement au moment de l'injection à travers la seringue pour leur sécrétion dans la cellule hôte (Marín et al., 2013). Il est probable que ces effecteurs n'acquièrent leur structure tridimensionnelle qu'au moment de l'interaction avec leur cible. Néanmoins, leur analyse structurale à travers l'extraction des segments bien prédits par Alphafold puis par leur alignement sur Foldseek (méthodologie établie par Teulet et al., 2023) a permis de prédire chez ces ET-

[146] Cette particularité est à relier à l'abondance de régions dites "en désordre" à l'intérieur des effecteurs (intrinsic disorder region).

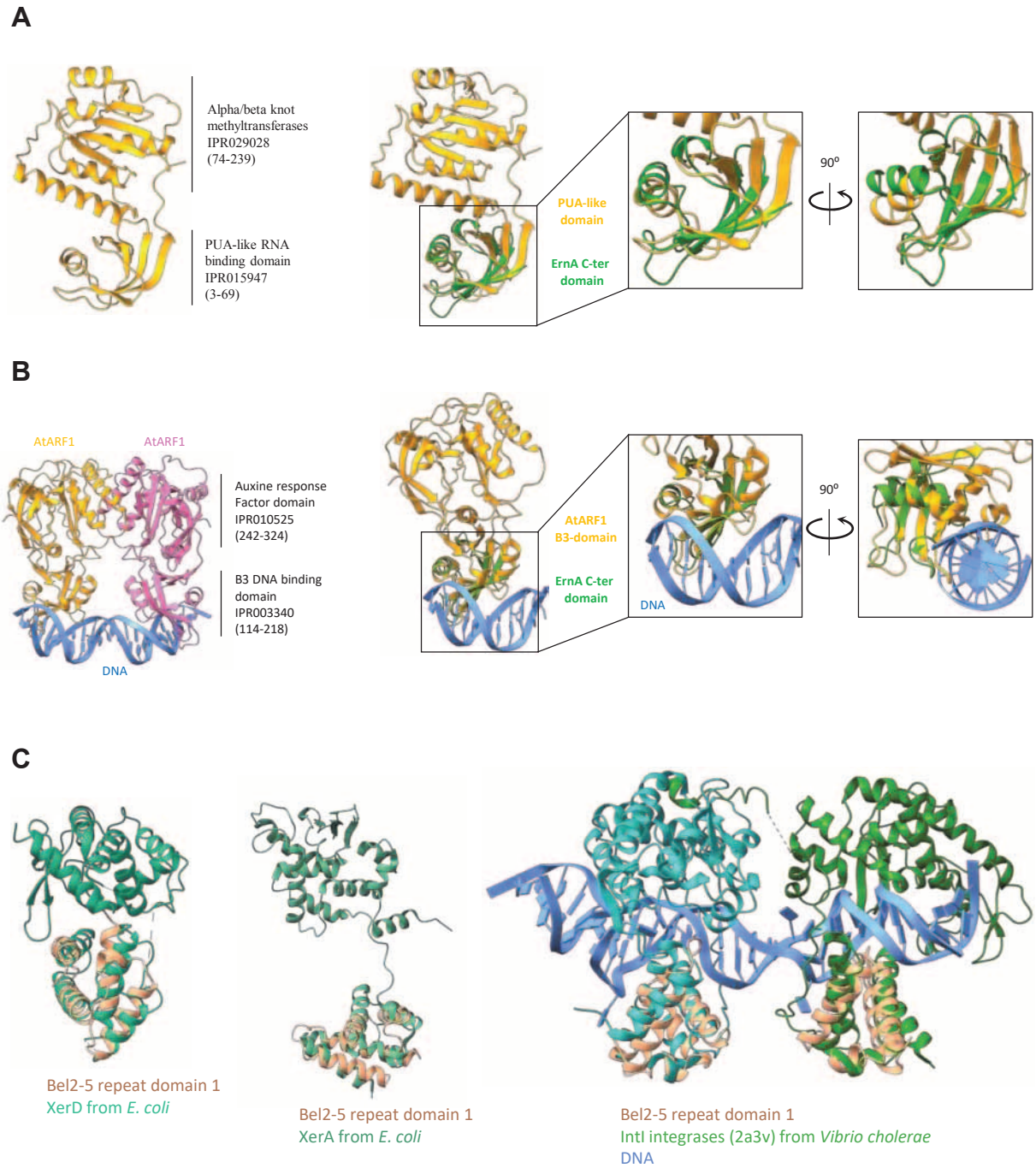


Figure 76. Fonction(s) putative(s) de divers domaines structuraux présents chez les ET-Nods (communication personnelle Albin Teulet et Sebastian Schornack).
L'extraction de la structure prédite de la partie C-ter d'ErnA pourrait correspondre au domaine de liaison à l'ARN (PUA-like) identifié chez des méthyltransférases (exemple d'une méthyltransférase identifiée chez la bactérie *Pseudomonas aeruginosa*) (**A**) ou au domaine B3 de liaison à l'ADN retrouvé chez les ARFs (exemple de AtARF1 identifié chez *Aradidopsis thaliana)* (**B**). Les numéros IPR correspondent à l'identifiant des domaines sur InterProScan. (**C**) Des domaines répétés sont retrouvés chez Bel2-5 et pourraient correspondre à des domaines CBD d'après leur prédiction structurale. Ces domaines sont retrouvés par exemples chez les recombinases XerD et XerA identifiées chez *Escherichia coli* (à gauche et au milieu) ou encore chez des intégrases telle qu'IntI retrouvée chez *Vibrio cholorae* (à droite). Ces domaines sont notamment connus pour interagir avec l'ADN. De telles structures ont également été identifiées chez Sup3 et d'autres SUMO-protéases telle que NopD.

Nods un certain nombre de domaines qui pourraient être à relier à leur fonction. Notamment sur la partie C-terminale d'ErnA, il peut être identifié un domaine qui pourrait correspondre à : i) un domaine PUA-like[147] souvent retrouvé chez des méthyltransférases[148], et ce plus particulièrement au niveau du domaine d'interaction avec l'ARN (figure 76A) ou ii) un domaine B3 de liaison à l'ADN, retrouvé notamment chez la famille des ARFs[149] (figure 76B) (Guilfoyle et al., 1998). Chez Sup3 et Bel2-5, des domaines répétés et conservés dans la première partie de la protéine ont pu être identifiés comme étant des CBD[150] (figure 76C) (communication personnelle Albin Teulet). Enfin, malgré le peu de points communs existants chez les ET-Nods, deux groupes semblent se former : ceux possédant un domaine SIM prédit (ErnA et Ubi1) et ceux possédant un domaine SUMO-protéase (Bel2-5 et Sup3).

3.2. Vers l'identification de nouveaux ET-Nods et de nouveaux processus Nod-indépendants

Les ET-Nods connus à ce jour n'ont été identifiés qu'en étudiant l'interaction symbiotique de souches de *Bradyrhizobium* sur deux légumineuses différentes : *A. indica* et *Glycine max*. L'analyse des propriétés symbiotiques sur *A. indica* de diverses souches de *Bradyrhizobium* séquencées a permis l'identification de cinq de ces six ET-Nods (ErnA, Sup3, Ubi1, Ubi2 et Ubi3) grâce à une analyse génomique comparative et une mutagénèse des gènes candidats (article 1). Actuellement, le groupe de Shin Okazaki développe ce même type d'approche de criblage d'une large diversité de souches de *Bradyrhizobium* sur le soja en utilisant la lignée mutée *nfr1* pour s'affranchir de l'activation de la nodulation par les FNs (communication personnelle, Shin Okazaki). On peut donc très bien penser que d'autres *Bradyrhizobium* différents d'USDA61 sont capables de noduler cette lignée mutante du soja en utilisant d'autres types d'ET-Nods différents de Bel2-5.

Pour aller plus loin dans l'identification de nouveaux ET-Nods, d'autres mutants de légumineuses affectés dans la perception des FNs pourraient être utilisés. Par ailleurs, ces lignées mutantes pourraient être utilisées pour "piéger" directement des rhizobiums à partir d'échantillons de sol inoculés. Il serait ainsi possible d'explorer encore plus la diversité des souches capables d'utiliser un processus Nod-indépendant, T3SS-dépendant ou pas, tout en s'affranchissant des étapes initiales d'isolement et de séquençage qui pourront être réalisées à posteriori à partir des nodules formés. Etant donné que jusqu'à présent, seules deux plantes ont été utilisées pour identifier des ET-Nods distincts à tra-

[147] Pseudo-Uridine synthase and Archaeosine transglycosase : initialement identifié chez les Archées, PUA est un motif de liaison à l'ARN conservé notamment chez des enzymes qui modifient post-transcriptionnellement l'ARNt et l'ARNr (InterProscan).

[148] Enzymes impliquées dans la méthylation de l'ADN ou des acides-aminés. La méthylation de l'ADN est notamment connue pour réprimer l'expression de gènes.

[149] Auxin response factors : famille de facteurs de transcriptions associés à une réponse précoce à l'auxine.

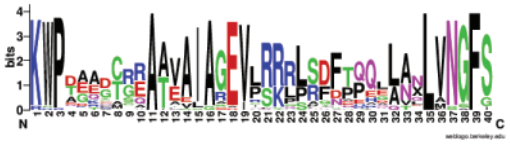
[150] Core-binding domain : domaine connu pour fixer l'ADN au niveau de motifs bien précis et retrouvé chez des recombinases par exemple.

DDA short repeat (36 aa)

DDA medium repeat (40 aa)

DDA long repeat (49 aa)

KWP repeat (40 aa)



SARA-DDA repeat (123 aa)

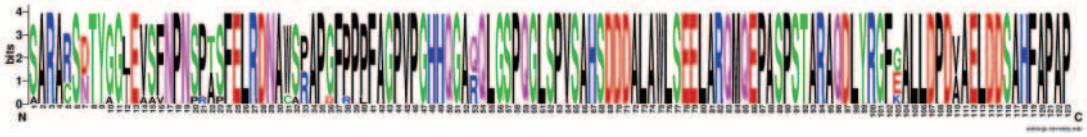


Figure 77. Identification de séquences répétées en tandem chez les SUMO-protéases identifiées chez *Bradyrhizobium*.
En plus des CBD, il peut être retrouvé des séquences répétées en tandem chez une majorité de SUMO-protéases (incluant Sup3 et Bel2-5) ont été identifiées à l'aide du logiciel en ligne RADAR (http://www.ebi.ac.uk/Tools/pfa/radar). Cinq catégories de séquences répétées en tandem peuvent être distinguées. On peut noter la présence d'acides-aminés conservés en début de séquence ("DDA", "KWP" ou "SARA"). Ces séquences répétées présentent pour la plupart le motif "DDA" mais des différences importantes peuvent être observées dans le nombre de répétions et la longueur ainsi que dans la conservation des acides-aminés. Weblogo réalisés en ligne sur https://weblogo.berkeley.edu/logo.cgi.

vers un nombre restreint de souches (limité par ailleurs à des *Bradyrhizobium*), on peut s'attendre à ce que de nouveaux ET-Nods soient découverts dans de nouveaux contextes symbiotiques. Il n'est pas à exclure non plus que d'autres processus Nod-indépendant, basés sur d'autres systèmes de sécrétions ou d'autres molécules signales, émergent dans le futur.

3.3. Les SUMO-protéases : un vivier d'ET-Nods ?

Les T3Es de la famille des SUMO-protéases sont les effecteurs les plus souvent retrouvés au sein des *Bradyrhizobium* (Teulet et al., 2020). Sur les 198 souches analysées dans l'article 1, 91 présentent un T3SS et parmi celles-ci uniquement, 228 SUMO-protéases ont pu être identifiées. On observe ainsi en moyenne plus de deux SUMO-protéases par souche possédant un T3SS, ce nombre pouvant monter jusqu'à 15 chez la souche *B. septentrionale* 1S1. La plupart de ces protéines présentent une architecture similaire, c'est-à-dire une très longue protéine (1200 aa en moyenne), un domaine SUMO-protéase localisé en C-terminal et une large variation de séquence sur tout le reste de la protéine avec régulièrement la présence de séquences répétées en tandem. La longueur de ces séquences répétées varie entre 36 et 123 aa et le nombre de répétition entre 2 et 20 (figure 77). Même si la longueur de ces séquences varie, elle rappelle celles observées chez les effecteurs TALs. Chez ces T3ES, la longueur des séquences répétées est souvent de 33 à 34 aa, elles sont fortement conservées entre chacune des répétitions (à l'exception du 12ème et 13ème aa) et leur nombre est bien souvent plus important (17 répétions en moyenne) (Mak et al., 2013; Perez-Quintero & Szurek, 2019). Il est donc difficile de penser que les séquences répétées en tandem identifiées chez nos SUMO-protéases puissent intervenir dans la reconnaissance de motifs spécifiques d'ADN à travers un mécanisme similaire à celui décrit chez les TALs. Par contre, il se pourrait qu'elles interviennent dans la reconnaissance de cibles protéiques spécifiques, plus ou moins à proximité de l'ADN.

L'analyse phylogénétique de cette famille de T3Es souligne une forte diversité de ces SUMO-protéases avec des clades fortement distincts les uns des autres et donc sûrement avec des propriétés et des rôles différents. On a pu montrer que les souches possédant un homologue de Sup3 sont généralement capables de noduler *A. indica* à l'inverse de celles présentant un homologue de Bel2-5. La spécificité de l'action de ces SUMO-protéases est bien illustrée dans l'article 1, dans lequel nous avons montré que la capacité de nodulation du mutant ORS3257Δ*ernA* pouvait être restaurée en introduisant l'homologue de *sup3* des souches de *Bradyrhizobium* sp. NAS96.2 et *B. archetypum* WSM1744. Par contre, aucune complémentation n'est obtenue en introduisant *bel2-5*, qui est pourtant responsable de l'activité de nodulation de la souche *B. elkanii* USD61 chez le soja : cela suggère une spécificité hôte-dépendante dans la capacité des ET-Nods à activer le programme de nodulation. Dans ce sens, il serait intéressant de voir si les souches possédant un homologue de Bel2-5 sont aussi

capables de noduler la lignée *nfr1* du soja comme le fait USDA61. On ne peut donc pas exclure que les souches de *Bradyrhizobium* présentant d'autres SUMO-protéases et appartenant à d'autres clades que Sup3 ou Bel2-5 puissent utiliser ces T3Es pour activer la nodulation chez d'autres espèces de légumineuses.

De nombreuses connaissances sont donc à acquérir pour mieux comprendre le mode d'action et la spécificité de cette famille de T3Es, qui apparaît être très importante dans le contrôle des interactions plantes-bactéries en général. En effet, les SUMO-protéases sont aussi souvent identifiées comme étant des déterminants clés pour promouvoir la pathogénicité des bactéries telle que *Xanthomonas* (Kim et al., 2013; Figaj et al., 2019). De manière surprenante, il a pu être montré que des SUMO-protéases telles que XopD chez *Xanthomonas* ou NopD chez *Bradyrhizobium* peuvent aussi être impliquées dans la déconjugaison de l'ubiquitine, qui est aussi une modification post-traductionnelle des protéines, ce qui suggère un rôle encore plus vaste de cette famille d'enzymes (Pruneda et al., 2016; Li et al., 2023). Le rôle symbiotique et le mode d'action de ces T3Es au sein du genre *Bradyrhizobium* reste donc une véritable énigme compte tenu de la diversité de cette famille et des multiples rôles qu'elle peut jouer. Cet aspect peut être illustré par la capacité de Bel2-5 à promouvoir ou compromettre la symbiose en fonction de la plante hôte (Ratu, Hirata, et al., 2021), qui est capable de déclencher la nodulation dans un contexte précis ou d'être un facteur d'incompatibilité dans l'autre.

4. Mode d'action des ET-Nods

4.1. Lien entre ET-Nods et partenaires SUMOylés

L'initiation de la caractérisation fonctionnelle des ET-Nods à travers le projet ANR "ET-Nod" a permis de mettre en lumière une possibilité d'interaction entre certains ET-Nods et des partenaires SUMOylés. Dans ce sens, une première expérimentation de Co-IP suggère qu'ErnA peut interagir *in vitro* avec au moins un des peptides SUMO (Ae07g02130) identifiés chez *A. evenia*, sans doute *via* son domaine SIM. Il est donc fort probable qu'ErnA puisse interagir de manière non-covalente avec une cible SUMOylée. Cette interaction pourrait modifier de nombreux aspects de la biologie des protéines SUMOylées et ciblées par ErnA, tels que leur structure, leur localisation mais aussi leurs fonctions qui conduiraient, par des mécanismes encore inconnus, à l'organogenèse nodulaire. Il est néanmoins nécessaire de confirmer ce résultat en utilisant entre-autres les mutants disponibles pour valider la fonctionnalité du SIM et d'autres types d'approches cette fois-ci *in vivo* (double-hybride, FRET-FLIM[151], luciférase fractionnée[152]). En parallèle, l'identification d'un domaine SIM chez Ubi1 suggére-

[151] Transfert d'énergie entre molécules fluorescentes (Fluorescence Resonance Energy Transfer).

rait que cet ET-Nod puisse agir de manière similaire à ErnA pour déclencher la nodulation. Il serait ainsi intéressant de confirmer sa fonctionnalité et sa capacité à lier des peptides SUMO.

L'existence de deux autres ET-Nods (Bel2-5 et Sup3) codant pour des SUMO-protéases vient renforcer cette hypothèse de régulation de certaines cibles modifiées post-traductionnellement par SUMOylation. L'activité de déSUMOylation de ces deux effecteurs a été prouvée être requise pour leur fonction symbiotique à travers la construction de formes mutées dans la triade catalytique du site actif du domaine SUMO-protéase, qui perdent alors la capacité de complémenter les mutants de délétion USDA61Δ*bel2-5* ou WSM1744Δ*sup3* (Ratu, Hirata, et al., 2021 et chapitre 3). De plus, l'activité protéase de Sup3 a été démontrée à travers sa capacité *in vitro* à cliver la liaison GST-SUMO. Dès lors, il peut être envisageable que la ou les cibles de ces SUMO-protéases soient les mêmes que celle(s) d'ErnA mais que dans ce cas, ce soit par déSUMOylation de ces dernières que l'organogenèse nodulaire puisse être initiée. Il pourrait être alors envisagé pour identifier ces cibles une comparaison du SUMOylome *d'A. evenia* en réponse ou pas à l'expression de *sup3*.

4.2. ErnA : un ET-Nod agissant sur des régulateurs centraux de la plante

Pour mieux comprendre le mode d'action des ET-Nods au niveau moléculaire, l'identification de leur(s) cible(s) est une priorité. L'émergence du modèle plante *A. evenia* N21 pour l'étude de cette symbiose nous permet aujourd'hui d'envisager de nombreuses approches fonctionnelles pour mieux comprendre le mode d'actions des ET-Nods chez les légumineuses (article 2). C'est pourquoi différentes approches ont été menées ou sont actuellement en cours pour déterminer dans un premier temps les cibles d'ErnA, le premier ET-Nod identifié chez *Aeschynomene*. La réalisation de deux analyses transcriptomiques comparatives chez *A. evenia* ainsi que l'obtention d'un potentiel interactome d'ErnA chez *N. benthamiana* a permis de mettre en évidence un certain nombre de candidats. Il apparaît qu'ErnA semble agir sur des régulateurs centraux puisqu'un ensemble de cascades de signalisations s'activent en réponse à l'expression d'*ernA*. En effet, un grand nombre de facteurs de transcriptions et de récepteurs/protéines kinases sont surexprimés en réponse à l'induction d'*ernA*. C'est toute une réaction en chaîne qui semble s'opérer et qui est particulièrement identifiable dans l'analyse pDEX conduite sur des lignées transgéniques d'*A. evenia* exprimant seulement *ernA*. De plus, il se pourrait qu'ErnA soit également impliqué dans la répression des mécanismes de défense de la plante en réponse à l'inoculation par la bactérie, comme suggéré dans l'article 2. L'établissement de l'interactome d'ErnA chez le tabac a également montré que des cibles impliquées dans les réponses transcriptionnelles de la plante pourraient interagir avec ErnA. Ces cibles potentielles sont au cœur de l'adaptation de la plante en réponse à des stress biotiques et abiotiques ; que

[152] Ou split-luciferase en anglais.

ce soit directement à travers des facteurs de transcriptions, des régulateurs épigénétiques ou des complexes protéiques pouvant réguler la transcription des gènes.

L'ensemble de ces observations est à relier avec la capacité d'ErnA à lier les acides nucléiques (Teulet et al., 2019). Il est alors fortement probable qu'en étant associé à l'ARN ou l'ADN, ErnA puisse réguler d'une manière qui nous est encore inconnue l'expression de gènes cibles pour déclencher le programme de nodulation. Cela pourrait par exemple s'opérer à travers un rôle épigénétique d'ErnA qui modifierait l'accessibilité de la chromatine. La SUMOylation a par ailleurs été montrée comme associée à l'hétérochromatine[153], à travers la SUMOylation des histones par exemple, ou aussi comme affectant la méthylation de l'ADN (Cubeñas-Potts & Matunis, 2013). ErnA étant sûrement capable de lier les peptides SUMO, il se pourrait qu'il interagisse avec des histones et promeuve ainsi l'expression de gènes spécifiques en les rendant plus accessibles à la machinerie de transcription. Une autre option pourrait être qu'ErnA agit au sein d'un complexe protéique constitué de multiples FTs qui favorise la transcription de certains gènes clés.

4.3. Des cibles candidates pour ErnA à explorer

Un certain nombre de candidats émergent des différentes analyses fonctionnelles effectuées (RNA-seq, Turbo-ID), mais nous discuterons ici de ceux qui nous paraissent les plus prometteurs et qui mériteraient qu'on aille plus loin pour les valider fonctionnellement.

Parmi la liste des FTs qui a été obtenue en couplant les analyses RNA-seq réalisées, le FT WRKY (Ae07g13440) est particulièrement intéressant. En effet, une récente revue met en évidence que les WRKYs ont un rôle beaucoup plus large que la modulation de la réponse immunitaire (Wang et al., 2023). Notamment, certains WRKYs régulent de nombreux processus développementaux et plus particulièrement, un lien entre WRKY et auxine stimulant le développement des racines latérales a pu être établi. Par exemple chez *Arabidopsis*, à travers certains ARFs, l'auxine contrôle l'expression d'*AtWRKY23* qui stimule le développement racinaire en modulant la distribution des auxines à travers les transporteurs PINs (Grunewald et al., 2012). Chez le blé, TaWRKY51 réprime la réponse hormonale de l'éthylène en faveur d'une signalisation auxinique qui promeut la formation de racines latérales (Hu et al., 2018). Il ne serait donc pas surprenant que le WRKY (Ae07g13440) qui a émergé lors de nos analyses puisse avoir un rôle développemental qui conduise à l'organogenèse du nodule ; le programme développemental d'une racine et d'un nodule étant intimement lié. Récemment, il a aussi été identifié chez *M. truncatula* la protéine MtWRP1[154], qui joue un rôle symbiotique notable dans la nodulation mais aussi dans la croissance racinaire (Chen et al., 2022). Un autre point notable

[153] Région de la chromatine connue pour être réprimée transcriptionnellement car difficile d'accès.
[154] WRKY-related protein.

observé sur notre WRKY (Ae07g13440) est son niveau d'expression dans le contexte Nod-indépendant, T3SS-indépendant, observé lors de l'interaction entre la souche photosynthétique ORS278 et *A. evenia* (Quilbé et al., 2021). En effet, sur les 91 WRKYs annotés dans le génome d'*A. evenia*, notre candidat se démarque car il est fortement exprimé de manière spécifique dans les nodules jeunes de 4 jours, avant de voir son niveau d'expression diminué progressivement lors de la maturation du nodule. Bien qu'obtenues dans un contexte symbiotique différent mais néanmoins Nod-indépendant, ces données d'expression semblent confirmer le rôle symbiotique que pourrait jouer ce FT.

Un autre candidat apparaît fort intéressant puisqu'il a déjà été décrit comme jouant un rôle symbiotique et code aussi pour un facteur de transcription : NF-YC (Ae06g34470). La recherche de cette famille de FTs dans le génome d'*A. evenia* a fait ressortir 8 gènes codants pour des NF-YC putatifs. De manière intéressante, une analyse phylogénétique de ces protéines avec les NF-YC identifiés chez *M. truncatula* montre que notre candidat semble atypique avec un domaine C-terminal plus long (communication Andreas Niebel). De plus, il apparaît aussi plus éloigné phylogénétiquement des NF-YC dont le rôle symbiotique chez *M. truncatula* a pu être montré (Baudin et al., 2015) ; suggérant que ce candidat est peut-être spécifique de la symbiose Nod-indépendante. Le complexe NF-Y, constitué de trois sous-unités (NF-YA, NF-YB et NF-YC), est connu pour son rôle symbiotique chez les légumineuses modèles à travers l'activation de gènes clés pour la nodulation et l'infection (Zanetti et al., 2017). Ces FTs sont dits "pionniers" car l'expression de ces gènes passerait par une ouverture de la chromatine à travers l'interaction de NF-YB et NF-YC avec des histones, facilitant ainsi l'action de NF-YA chez *M. truncatula* (communication Andreas Niebel).

Le marquage de proximité des interactants d'ErnA chez le tabac a mis en exergue un nombre important des sous-unités constituant le Médiateur ; faisant de ce complexe un candidat qui mérite notre attention. Impliqué dans le recrutement de l'ARN polymérase II, le Médiateur peut cibler spécifiquement des régions de l'ADN pour activer la transcription de gènes. Cette spécificité passe notamment par son interaction avec des FTs (Yang et al., 2016). A ce jour, le seul lien qui a pu être décrit entre la symbiose rhizobium-légumineuse et le complexe Médiateur passe par le gène *LjLAN*[155] qui code pour une protéine homologue à une putative sous-unité du complexe AtMED2 décrite chez *A. thaliana* (Suzaki et al., 2019). LjLAN pourrait agir au sein d'une voie symbiotique parallèle à celle déjà décrite et a été montré requis pour l'infection intracellulaire ainsi que pour l'adaptation de la bactérie au sein du nodule, soulignant le rôle symbiotique du Médiateur chez le lotier. Par ailleurs, il a été proposé que l'auxine recrute des ARFs pour réguler l'expression du FT LBD16 en dégradant des répresseurs

[155] LACK OF SYMBIONT ACCOMODATION.

transcriptionnels (IAA14[156], TPL[157]) qui forment ensemble un complexe avec le Médiateur chez *A. thaliana* (Ito et al., 2016). Le FT LB16 est connu pour réguler la formation de racines latérales mais aussi l'organogenèse du nodule à travers NIN et l'auxine (Schiessl et al., 2019). Il serait alors envisageable qu'ErnA, dans un contexte symbiotique, puisse faire partie d'un tel complexe pour activer le programme de nodulation. L'identification par Turbo-ID chez le tabac du FT nommé TPL parmi les interactants possibles d'ErnA vient appuyer cette hypothèse et suggère là encore la formation d'un complexe protéique pour réguler le processus d'organogenèse nodulaire.

Enfin, dans la littérature, plusieurs exemples peuvent être cités pour souligner le lien possible entre ces candidats et la SUMOylation, comme cela a pu être suggéré précédemment avec les ET-Nods. Il a ainsi été montré que la SUMOylation de NF-YC aide à l'assemblage du complexe NF-Y chez *A. thaliana* et permet une meilleure réponse transcriptomique des gènes cibles de la plante (Huang et al., 2023). De même, un WRKY a été montré comme pouvant être SUMOylé pour moduler l'immunité de la plante à travers l'interaction de certaines MAPK chez *A. thaliana via* un motif SIM (Verma et al., 2021). Dans ce contexte, la SUMOylation permet de réguler l'activité transcriptionnelle de AtWRKY33 à un stade précoce après infection avec un pathogène. La SUMOylation des FTs impliqués dans la signalisation auxinique est également souvent retrouvée pour moduler de nombreux processus développementaux (Roy & Sadanandom, 2021). De manière plus globale, une vaste diversité de FTs sont connus pour être régulés par SUMOylation/déSUMOylation, ce qui peut modifier leur capacité de liaison à l'ADN, leur accessibilité à la chromatine, leur abondance et leur localisation (Roy & Sadanandom, 2021). Mises bout à bout, ces données viennent renforcer l'idée qu'ErnA peut interagir avec une cible SUMOylée à travers son motif SIM et qu'un "hub" protéique pouvant rassembler plusieurs ET-Nods et FTs est mis en place pour intensifier la réponse moléculaire responsable de l'induction de la nodulation.

Des approches sont maintenant à développer pour confirmer ou rejeter ces candidats. Pour les cibles protéiques identifiées par Turbo-ID chez le tabac, il faudra dans un premier temps les confirmer chez *A. evenia* en utilisant la même stratégie. Des expérimentations de double-hybride en levure, des RT-qPCR ainsi que des approches *in vitro* par Co-IP ou directement *in planta* par BiFC[158] ou FRET-FLIM sont envisagées pour valider ces candidats. Du ChIP-Seq peut également être réalisé pour confirmer la liaison d'ErnA à une région spécifique de l'ADN. La validation des cibles d'ErnA est prévue dans le cadre du projet ANR "ET-Nod".

[156] Indole 3-acetic acid 14. Les IAAs sont connus pour interagir avec les ARFs et ainsi inhiber la transcription de gènes normalement activée par les ARFs.
[157] TOPLESS, corépresseur transcriptionnel.
[158] Bimolecular fluorescence complementation.

5. Décryptage de la voie de signalisation empruntée par la symbiose T3SS-dépendante

Pour déchiffrer les déterminants génétiques impliqués dans le processus Nod-indépendant, T3SS-dépendant, nous avons mis à profit les mutants EMS disponibles dans les gènes requis pour la symbiose Nod-indépendante, T3SS-indépendante chez *A. evenia* lignée 76. Six gènes ont donc été testés (*AeCRK*, *AePOLLUX*, *AeCCaMK*, *AeCYCLOPS*, *AeNSP2* et *AeNIN*) pour déterminer s'ils sont aussi requis lors de la nodulation T3SS-dépendante déclenchée par diverses souches de *Bradyrhizobium* présentant différents effectomes. Les travaux réalisés dans l'article 2 de cette thèse montrent ainsi que la voie de signalisation empruntée par la symbiose Nod-indépendante, T3SS-dépendante est conservée partiellement avec celle des autres processus symbiotiques. Par ailleurs, nous avons montré qu'AePOLLUX, AeCCaMK et AeCYCLOPS peuvent être dispensables selon la souche de *Bradyrhizobium* inoculée ; mais que dans tous les cas, AeNSP2 et AeNIN semblent être requis (figure 78). Il apparaît donc que les T3Es peuvent intervenir à différents niveaux dans la voie de signalisation. Il peut également être possible qu'une voie parallèle à celle déjà décrite puisse être activée par les ET-Nods. Dans cette optique, un facteur X qui reste à déterminer pourrait faire le lien entre ces deux voies parallèles et pourrait se situer entre CCaMK et CYCLOPS (Limpens & Bisseling, 2014) ; tout comme cela a été proposé pour LjLAN (Suzaki et al., 2019).

La capacité des souches à court-circuiter un ou plusieurs déterminants clés apparaît être liée à leur contenu en ET-Nods. Les capacités symbiotiques de la souche *Bradyrhizobium* sp. ORS86 qui possède deux ET-Nods (ErnA et Sup3) vient appuyer cette hypothèse. En effet, la souche sauvage est capable de noduler l'ensemble des mutants à l'exception de *nsp2* et *nin* alors que la mutation d'un seul des deux ET-Nods abolit la nodulation auparavant observée sur les mutants *ccamk* et *cyclops*. Il semble donc que la présence d'un seul ET-Nod ne soit pas suffisante pour court-circuiter une partie des déterminants symbiotiques. La voie de signalisation Nod-indépendant, T3SS-dépendant pourrait ainsi impliquer dans certains cas une coopérativité entre les ET-Nods pour déclencher le programme de nodulation. Les mécanismes par lesquels ces ET-Nods coopèrent demeurent méconnus mais on peut imaginer une ou des cibles communes SUMOylées dont l'activité serait modulée par ces différents T3Es.

Pour aller plus loin dans la caractérisation de la voie de signalisation T3SS-dépendante, il pourrait être envisagé de réaliser une mutagénèse EMS similaire à celle qui a été faite sur *A. evenia* N76, mais sur la lignée modèle de la symbiose T3SS-dépendante : *A. evenia* N21. Le criblage des mutants ayant perdu leur capacité de nodulation en réponse à l'inoculation d'une souche au phénotype symbiotique fort (telle qu'ORS86), mènerait à une meilleure compréhension de la voie de signalisation qui est recrutée. Ce type d'expérimentation demeure cependant excessivement lourd au niveau humain

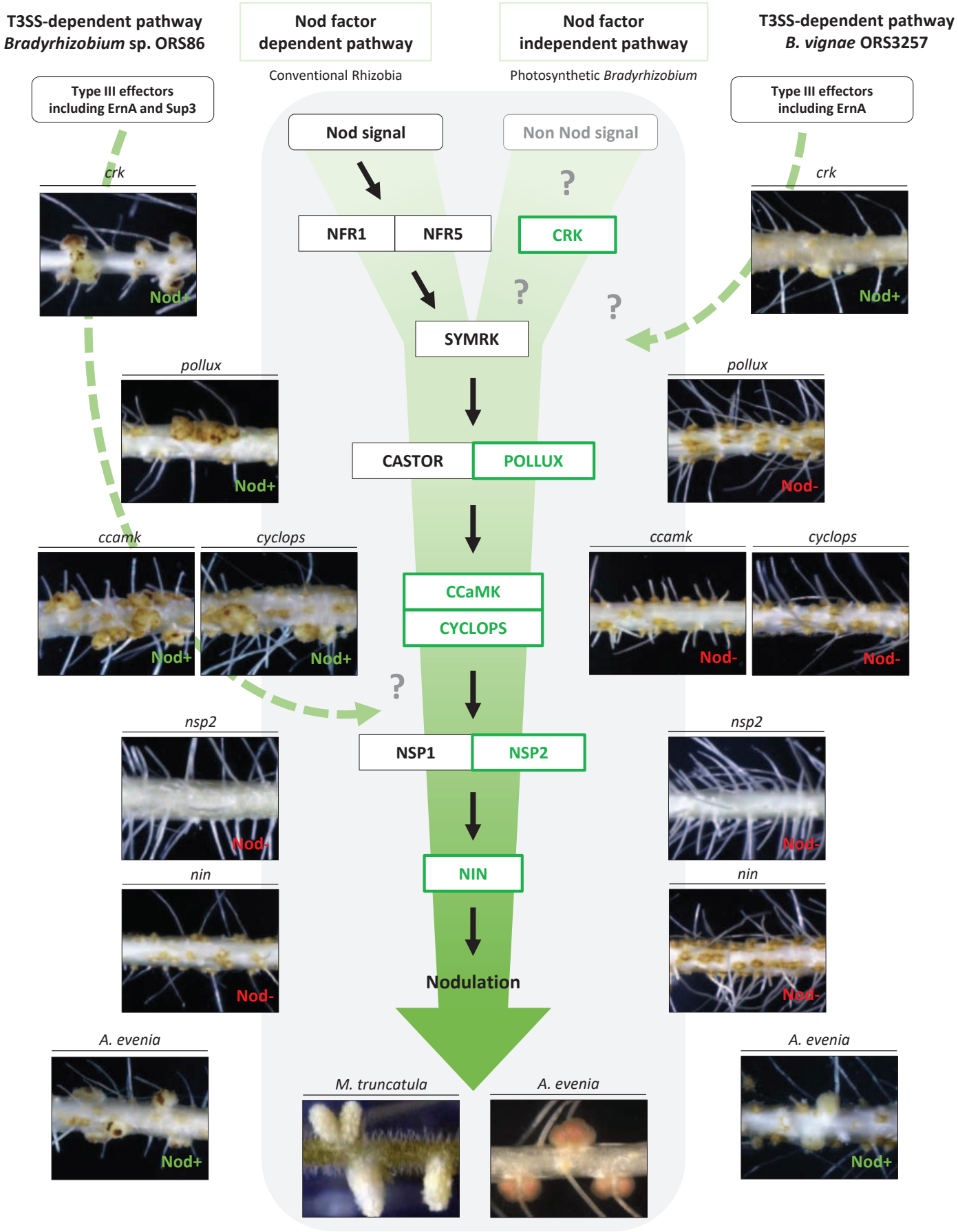


Figure 78. La symbiose Nod-indépendante, T3SS dépendante recrute divers acteurs de la voie de signalisation symbiotique commune qui diffèrent selon les souches et les ET-Nods injectés.
Les déterminants symbiotiques dont les mutants sont disponibles chez *A. evenia* N76 apparaissent en vert dans la voie. Deux modèles émergent selon les souches inoculées pour mieux décrypter la symbiose T3SS-dépendante chez les *Bradyrhizobium* non-photosynthétique. Ainsi, les T3Es pourraient intervenir à différents niveaux, dans la voie de signalisation (indiquée par les flèches vertes en pointillés). Les phénotypes observés pour chaque mutant sont notés en rouge pour Nod- et en vert pour Nod+. Le point d'interrogation indique qu'à ce jour le mécanisme de signalisation mis en place demeure inconnu.

et financier et il est donc important de prendre en compte ces critères avant de se lancer dans une nouvelle approche de mutagénèse EMS.

6. Futurs défis et perspectives

Le contexte démographique actuel (plus de 8 milliards d'êtres humains sur Terre) ainsi que l'omniprésence des conséquences du réchauffement climatique sur notre planète (hausse des températures, augmentation des catastrophes naturelles, perte de biodiversité, pénuries alimentaires, etc.) constituent des défis à relever pour envisager un avenir plus serein pour l'humanité et notre planète.

La culture des légumineuses, à travers la symbiose fixatrice d'azote, ne requiert pas d'engrais azotés pour maintenir un bon rendement. En plus de cet intérêt agronomique et environnemental, les légumineuses sont aussi une source de protéines pour l'alimentation, faisant d'elles une alternative à la consommation de protéines d'origine animale qui nécessite de nombreuses ressources et présente ainsi un coût écologique énorme. La recherche sur la symbiose rhizobium-légumineuse présente donc un réel intérêt pour développer une agriculture plus durable et respectueuse de l'environnement tout en répondant aux besoins de notre planète.

6.1. Améliorer les performances symbiotiques des rhizobiums *via* les ET-Nods dans un contexte Nod-dépendant

Les souches utilisant un processus symbiotique Nod-indépendant, T3SS-dépendant possèdent pour la majorité à la fois les gènes *nod* et ceux codant pour le T3SS. Cela suggère que les gènes *nod*, bien que dispensables dans la symbiose avec *A. indica*, jouent sûrement un rôle majeur chez les plantes hôtes originaires de ces souches. En effet, le système d'étude utilisé pour cette symbiose est artificiel et abroge la signalisation des FNs où seul l'effet des ET-Nods déclenchant la nodulation peut être observé. On peut alors se demander quel est l'intérêt pour ces souches de posséder à la fois des ET-Nods et des FNs dans leur contexte symbiotique d'origine, et si un effet cumulatif des deux pourrait améliorer l'efficience symbiotique ?

Initialement isolée du niébé (*Vigna unguiculata*), la souche *B. vignae* ORS3257 utilise ses FNs pour déclencher la nodulation chez son hôte naturel (Songwattana et al., 2021). Pour mieux comprendre l'intérêt de cette souche de posséder un ET-Nod (ErnA), une collaboration avec le groupe de Neung Teaumroong (Université de Suranaree, Thaïlande) a conduit à la publication d'un article montrant l'importance du rôle du T3SS selon les espèces de *Vigna* considérées (annexe 1). Sur *V. radiata*, le T3SS d'ORS3257 rend incompatible l'interaction *via* la sécrétion de NopP qui induit une HR ; alors que sur *V. mungo* ou *V. unguiculata*, le T3SS joue un rôle bénéfique très important dans la capacité

d'infection et le nombre de nodules induits à travers au moins deux effecteurs : NopT et NopAB. Par contre, sur les trois espèces de *Vigna* testées, aucun rôle notable d'ErnA n'a pu être observé, indiquant que dans ces cas-là, l'induction de la nodulation est principalement sous le contrôle des FNs. On ne peut cependant pas exclure que dans d'autres contextes impliquant d'autres plantes hôtes, possédant par exemple des récepteurs LysM-RLK moins efficients dans la reconnaissance des FNs, ou que dans d'autres conditions environnementales peu favorables avec un inoculum bactérien plus faible ou en compétition avec une microflore indigène, les ET-Nods contribueraient significativement à la mise en place d'une symbiose efficiente tout en permettant d'élargir le spectre d'hôte des bactéries. Il serait donc intéressant d'examiner dans un premier temps, à travers la collection de mutants présents dans différents gènes clés (*T3SS*, *ET-Nod*, *nod*) chez différentes souches de *Bradyrhizobium*, le rôle que peuvent-jouer les ET-Nods chez différentes plantes hôtes où les FNs sont supposés être le signal clé susceptible de déclencher la nodulation.

Par ailleurs, l'émergence d'un nombre considérable de souches de *Bradyrhizobium* isolées de la rhizosphère du riz et capables de noduler *A. indica* (article 3) soulève la question de l'intérêt pour le riz d'avoir de telles souches dans son environnement proche. Il a déjà été montré que certains *Bradyrhizobium* photosynthétiques améliorent la croissance et le rendement chez un riz africain (Chaintreuil et al., 2000). Il serait donc important de mieux comprendre le rôle de ces souches (photosynthétiques ou non) lors de leur interaction avec différentes variétés de riz en réalisant des essais à la fois en laboratoire et au champ et en identifiant les déterminants génétiques impliqués dans cette interaction. Cette nouvelle diversité disponible pourrait dès lors permettre le développement de bio-inoculants permettant d'améliorer le rendement de non-légumineuses (effets multiples PGPR).

6.2. Vers un transfert de la capacité symbiotique aux céréales ?

Il a été proposé depuis plusieurs décennies de transférer la capacité symbiotique aux céréales pour limiter l'utilisation d'engrais azotés, qui sont associés à une pollution élevée et à un coût de production et d'achat conséquent. C'est tout un consortium international de scientifiques faisant partie du projet ENSA[159] qui travaille sur cette thématique, de manière à augmenter durablement les rendements de production des céréales, et ce plus particulièrement en Afrique. Dans cette optique, il est nécessaire d'avoir au sein des céréales le programme symbiotique caractérisé chez les légumineuses modèles pour espérer leur permettre de s'associer aux rhizobiums. Le fait qu'elles soient capables d'être mycorhizées *via* une voie de signalisation en partie commune à celle de la nodulation est très encourageant (An et al., 2010; Nakagawa & Imaizumi-Anraku, 2015). Cependant, cette signalisation symbiotique est très sophistiquée et cet objectif constitue un réel défi pour la communauté scienti-

[159] Enabling Nutrient Symbioses in Agriculture (https://www.ensa.ac.uk/).

fique. Des avancées prometteuses ont été récemment obtenues chez le lotier, où il a été montré que les homologues de LjNFR1 et LjNFR5 identifiés chez l'orge sont capables de complémenter une lignée mutée dans ces récepteurs aux FNs (Rübsam et al., 2023). Malgré une divergence évolutive entre ces récepteurs il y a plus de 200 Ma, il apparaît que ceux de l'orge possèdent toujours cette capacité à déclencher l'organogenèse nodulaire chez *L. japonicus*. On peut donc supposer que d'autres déterminants génétiques de la nodulation aient pu être conservés chez les céréales, bien que ces deux familles (légumineuses et graminées) soient aujourd'hui très divergentes. De nouvelles connaissances sur les mécanismes moléculaires impliqués dans la stabilisation et la polarité du cordon d'infection ont également été acquises : elles permettent d'améliorer notre vision sur les acteurs requis pour l'infection intracellulaire par les rhizobiums et d'ainsi essayer de l'adapter aux céréales (Lace et al., 2023; Su et al., 2023). Cependant, la synchronisation entre infection et organogenèse est requise pour avoir un nodule infecté intracellulairement, ce qui représente un autre défi à relever. Par la suite, d'autres étapes seront encore nécessaires pour réussir à transférer cette capacité symbiotique aux céréales, dont notamment le maintien des bactéries dans le nodule pour leur permettre de fixer l'azote qui devra ensuite être transmis à la plante. Les mécanismes moléculaires des étapes plus tardives de la nodulation (fixation d'azote et infection chronique des bactéroïdes dans le symbiosome) sont encore peu connus et constituent des lacunes importantes de connaissances à combler avant de pouvoir imaginer développer des non-légumineuses interagissant avec des bactéries fixatrices d'azote.

6.3. Les ET-Nods : un moyen de faciliter la nodulation chez les non-légumineuses ?

A travers la sécrétion d'effecteurs spécifiques, la symbiose Nod-indépendant, T3SS-dépendant s'affranchit de certains déterminants génétiques clés de la voie symbiotique décrite chez les légumineuses modèles. Il serait alors possible que l'utilisation d'ET-Nods permette de faciliter l'activation du programme de nodulation chez les non-légumineuses. Il pourrait dans ce sens être intéressant de tester le rôle des ET-Nods une fois injectés dans les cellules de céréales et d'observer les réponses développementales qui pourraient être induites. Ces effecteurs permettraient-ils à eux seuls d'induire des pseudo-nodules chez les céréales ? Si c'est le cas, cela suggérerait un rôle bien plus vaste des ET-Nods que ceux décrits chez les légumineuses, qui faciliterait le processus d'organogenèse dans un contexte non symbiotique.

Par ailleurs, les recherches actuelles sur le transfert de la capacité symbiotique aux céréales se focalisent essentiellement sur les déterminants plantes. Il serait intéressant de travailler en parallèle sur les déterminants bactériens pour faciliter ces recherches. Cela pourrait par exemple être envisagé par le transfert chez un rhizobium d'intérêt des déterminants requis pour initier l'infection et

l'organogenèse, ce qui passerait sûrement en partie par un cocktail précis d'effecteurs contenant entre autres des ET-Nods et des T3Es réprimant la réponse immunitaire de la plante pour permettre l'acceptation intracellulaire de la bactérie. De plus, une infection initialement intercellulaire par crack-entry pourrait faciliter le processus pour obtenir un nodule infecté. En comparaison, de nombreux acteurs sont requis pour l'élaboration et le fonctionnement d'un cordon d'infection.

C'est pour l'ensemble de ces raisons que l'étude de ces processus symbiotiques alternatifs qui paraissent atypiques demeure une source exploratoire importante pour mieux répondre aux enjeux climatiques et démographiques de notre planète. Mieux les comprendre conduira ainsi à une meilleure vision globale de l'apparition et du fonctionnement des symbioses fixatrices d'azote sur Terre et peut aider à faire émerger d'autres voies à explorer pour répondre à ces défis sociétaux et environnementaux.

Références bibliographiques

Abascal, E., Gómez-Coma, L., Ortiz, I., & Ortiz, A. (2022). Global diagnosis of nitrate pollution in groundwater and review of removal technologies. *Science of the Total Environment*, *810*. https://doi.org/10.1016/j.scitotenv.2021.152233

Abreu, I., Saéz, Á., Castro-Rodríguez, R., Escudero, V., Rodríguez-Haas, B., Senovilla, M., Larue, C., Grolimund, D., Tejada-Jiménez, M., Imperial, J., & González-Guerrero, M. (2017). *Medicago truncatula* zinc-iron permease6 provides zinc to rhizobia-infected nodule cells. *Plant Cell and Environment*, *40*(11), 2706–2719. https://doi.org/10.1111/pce.13035

Adams, D. G., Bergman, B., Nierzwicki-Bauer, S. A., Duggan, P. S., Rai, A. N., & Schüßler, A. (2013). Cyanobacterial-Plant Symbioses. In *The Prokaryotes: Prokaryotic Biology and Symbiotic Associations*. https://doi.org/10.1007/978-3-642-30194-0

Ahemad, M. (2014). Mechanisms and applications of plant growth promoting rhizobacteria : Current perspective. *Journal of King Saud University - Science*, *26*(1), 1–20. https://doi.org/10.1016/j.jksus.2013.05.001

Ahemad, M., & Khan, M. S. (2013). Pesticides as Antagonists of Rhizobia and the Legume-Rhizobium Symbiosis: a Paradigmatic and Mechanistic Outlook. *Biochemistry & Molecular Biology*, *1*(4), 63. https://doi.org/10.12966/bmb.12.02.2013

Alazard, D. (1985). Stem and Root Nodulation in *Aeschynomene* spp . *Applied and Environmental Microbiology*, *50*(3), 732–734. https://doi.org/10.1128/aem.50.3.732-734.1985

Alazard, Didier. (1990). Nitrogen fixation in pure culture by rhizobia isolated from stem nodules of tropical Aeschynomene species. *FEMS Microbiology Letters*, *68*(1–2), 177–182. https://doi.org/10.1016/0378-1097(90)90147-I

Alexander, C., & Rietschel, E. T. (2001). Bacterial lipopolysaccharides and innate immunity. *Journal of Endotoxin Research*, *7*(3), 167–202. https://doi.org/10.1179/096805101101532675

Allen, B. L., & Taatjes, D. J. (2015). The Mediator complex: A central integrator of transcription. *Nature Reviews Molecular Cell Biology*, *16*(3), 155–166. https://doi.org/10.1038/nrm3951

Amadou, C., Pascal, G., Mangenot, S., Glew, M., Bontemps, C., Capela, D., Carrère, S., Cruveiller, S., Dossat, C., Lajus, A., Marchetti, M., Poinsot, V., Rouy, Z., Servin, B., Saad, M., Schenowitz, C., Barbe, V., Batut, J., Médigue, C., & Masson-Boivin, C. (2008). Genome sequence of the β-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Research*, *18*(9), 1472–1483. https://doi.org/10.1101/gr.076448.108

Amor, B. Ben, Shaw, S. L., Oldroyd, G. E. D., Maillet, F., Penmetsa, R. V., Cook, D., Long, S. R., Dénarié, J., & Gough, C. (2003). The *NFP* locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. *Plant Journal*, *34*(4), 495–506. https://doi.org/10.1046/j.1365-313X.2003.01743.x

An, G. H., Kobayashi, S., Enoki, H., Sonobe, K., Muraki, M., Karasawa, T., & Ezawa, T. (2010). How does arbuscular mycorrhizal colonization vary with host plant genotype? An example based on maize (*Zea mays*) germplasms. *Plant and Soil*, *327*(1), 441–453. https://doi.org/10.1007/s11104-009-0073-3

Andrews, M., & Andrews, M. E. (2017). Specificity in legume-rhizobia symbioses. *International Journal of Molecular Sciences*, *18*(4). https://doi.org/10.3390/ijms18040705

Antolín-Llovera, M., Ried, M. K., & Parniske, M. (2014). Cleavage of the symbiosis receptor-like kinase ectodomain promotes complex formation with nod factor receptor 5. *Current Biology*, *24*(4), 422–427. https://doi.org/10.1016/j.cub.2013.12.053

Arrighi, J.-F., & Cartieaux, F. (2015). Out of Water of a New Model Legume: The Nod-independent *Aeschynomene evenia*. *Biological Nitrogen Fixation*, *1–2*(July), 447–454. https://doi.org/10.1002/9781119053095.ch45

Arrighi, J.-F., Cartieaux, F., Brown, S. C., Rodier-goud, M., Boursot, M., Fardoux, J., Patrel, D., Gully, D., Fabre, S., Chaintreuil, C., & Giraud, E. (2012). *Aeschynomene evenia*, a model plant for studying the molecular

genetics of the Nod-independent rhizobium-legume symbiosis. *Molecular Plant-Microbe Interactions : MPMI*, *25*(7), 851–861. https://doi.org/http://dx.doi.org/10.1094/MPMI -02-12-0045-TA

Arrighi, J. F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L. C., Mirabella, R., De Carvalho-Niebel, F., Journet, E. P., Ghérardi, M., Huguet, T., Geurts, R., Dénarié, J., Rougé, P., & Gough, C. (2006). The *Medicago truncatula* lysine motif-receptor-like kinase gene family includes *NFP* and new nodule-expressed genes. *Plant Physiology*, *142*(1), 265–279. https://doi.org/10.1104/pp.106.084657

Avontuur, J. R., Palmer, M., Beukes, C. W., Chan, W. Y., Coetzee, M. P. A., Blom, J., Stępkowski, T., Kyrpides, N. C., Woyke, T., Shapiro, N., Whitman, W. B., Venter, S. N., & Steenkamp, E. T. (2019). Genome-informed *Bradyrhizobium* taxonomy: where to from here? *Systematic and Applied Microbiology*, *42*(4), 427–439. https://doi.org/10.1016/j.syapm.2019.03.006

Avontuur, J. R., Wilken, P. M., Palmer, M., Coetzee, M. P. A., Stępkowski, T., Venter, S. N., & Steenkamp, E. T. (2023). Complex evolutionary history of photosynthesis in *Bradyrhizobium*. *Microbial Genomics*, *9*(9), 1–13. https://doi.org/10.1099/mgen.0.001105

Azarakhsh, M., & Lebedeva, M. A. (2023). Lateral Root versus Nodule : The Auxin - Cytokinin Interplay. *Journal of Plant Growth Regulation*, *March*. https://doi.org/10.1007/s00344-023-10983-4

Baier, M. C., Barsch, A., Küster, H., & Hohnjec, N. (2007). Antisense repression of the *Medicago truncatula* nodule-enhanced sucrose synthase leads to a handicapped nitrogen fixation mirrored by specific alterations in the symbiotic transcriptome and metabolome. *Plant Physiology*, *145*(4), 1600–1618. https://doi.org/10.1104/pp.107.106955

Balint-Kurti, P. (2019). The plant hypersensitive response: concepts, control and consequences. *Molecular Plant Pathology*, *20*(8), 1163–1178. https://doi.org/10.1111/mpp.12821

Balloon, A. J., Maeda, J., Forshey, K., & Os, D. Den. (2015). Correction for Venkateshwaran et al., A role for the mevalonate pathway in early plant symbiotic signaling. *Proceedings of the National Academy of Sciences*, *112*(38), 5378. https://doi.org/10.1073/pnas.1516711112

Barnett, M. J., Fisher, R. F., Jones, T., Komp, C., Abola, A. P., Barloy-Hubler, F., Bowser, L., Capela, D., Galibert, F., Gouzy, J., Gurjal, M., Hong, A., Huizar, L., Hyman, R. W., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Palm, C., … Long, S. R. (2001). Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(17), 9883–9888. https://doi.org/10.1073/pnas.161294798

Barrera, L. L., Trujillo, M. E., Goodfellow, M., García, F. J., Hernández-Lucas, I., Dávila, G., Van Berkum, P., & Martínez-Romero, E. (1997). Biodiversity of bradyrhizobia nodulating *Lupinus* spp. *International Journal of Systematic Bacteriology*, *47*(4), 1086–1091. https://doi.org/10.1099/00207713-47-4-1086

Baudin, M., Laloum, T., Lepage, A., Rípodas, C., Ariel, F., Frances, L., Crespi, M., Gamas, P., Blanco, F. A., Zanetti, M. E., de Carvalho-Niebel, F., & Niebel, A. (2015). A phylogenetically conserved group of nuclear factor-Y transcription factors interact to control nodulation in legumes. *Plant Physiology*, *169*(4), 2761–2773. https://doi.org/10.1104/pp.15.01144

Belin, B. J., Busset, N., Giraud, E., Molinaro, A., Silipo, A., & Newman, Di. K. (2018). Hopanoid lipids: From membranes to plant-bacteria interactions. *Nature Reviews Microbiology*, *16*(5), 304–315. https://doi.org/10.1038/nrmicro.2017.173

Bender, G. L. (1988). The *nodD1* Gene from *Rhizobium* Strain NGR234 Is a Key Determinant in the Extension of Host Range to the Nonlegume *Parasponia* . In *Molecular Plant-Microbe Interactions* (Vol. 1, Issue 7, p. 259). https://doi.org/10.1094/mpmi-1-259

Benková, E., & Bielach, A. (2010). Lateral root organogenesis - from cell to organ. *Current Opinion in Plant Biology*, *13*(6), 677–683. https://doi.org/10.1016/j.pbi.2010.09.006

Boivin, C., Ndoye, I., Molouba, F., De Lajudie, P., Dupuy, N., & Dreyfus, B. (1997). Stem Nodulation in Legumes: Diversity, Mechanisms, and Unusual Characteristics. *Critical Reviews in Plant Sciences*, *16*(1), 1–30. https://doi.org/10.1080/07352689709701944

Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., Goormachtig, S., & Giraud, E. (2011). Nodulation of *Aeschynomene afraspera* and *A. indica* by photosynthetic *Bradyrhizobium* sp. strain ORS285: the Nod-dependent versus the Nod-independent symbiotic interation. *Molecular Plant-Microbe*

Interactions, *24*(11), 1359–1371.

Bonaldi, K., Gourion, B., Fardoux, J., Hannibal, L., Cartieaux, F., Boursot, M., Vallenet, D., Chaintreuil, C., Prin, Y., Nouwen, N., & Giraud, E. (2010). Large-scale transposon mutagenesis of photosynthetic *Bradyrhizobium* sp. strain ORS278 reveals new genetic loci putatively important for nod-independent symbiosis with *Aeschynomene indica*. *Molecular Plant-Microbe Interactions*, *23*(6), 760–770. https://doi.org/10.1094/MPMI-23-6-0760

Boogerd, F. C., & Van Rossum, D. (1997). Nodulation of groundnut by *Bradyrhizobium*: A simple infection process by crack entry. *FEMS Microbiology Reviews*, *21*(1), 5–27. https://doi.org/10.1016/S0168-6445(97)00004-1

Booth, N. J., Smith, P. M. C., Ramesh, S. A., & Day, D. A. (2021). Malate transport and metabolism in nitrogen-fixing legume nodules. *Molecules*, *26*(22), 1–13. https://doi.org/10.3390/molecules26226876

Borghi, L. (2010). Inducible Gene Expression Systems for Plants. *Methods in Molecular Biology*, *655*(June 2010), 65–75. https://doi.org/10.1007/978-1-60761-765-5

Bourke, P. M., Evers, J. B., Bijma, P., van Apeldoorn, D. F., Smulders, M. J. M., Kuyper, T. W., Mommer, L., & Bonnema, G. (2021). Breeding Beyond Monoculture: Putting the "Intercrop" Into Crops. *Frontiers in Plant Science*, *12*(November). https://doi.org/10.3389/fpls.2021.734167

Boutrot, F., & Zipfel, C. (2017). Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. *Annual Review of Phytopathology*, *55*, 257–286. https://doi.org/10.1146/annurev-phyto-080614-120106

Bozsoki, Z., Cheng, J., Feng, F., Gysel, K., Vinther, M., Andersen, K. R., Oldroyd, G., Blaise, M., Radutoiu, S., & Stougaard, J. (2017). Receptor-mediated chitin perception in legume roots is functionally separable from Nod factor perception. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(38), E8118–E8127. https://doi.org/10.1073/pnas.1706795114

Breakspear, A., Liu, C., Roy, S., Stacey, N., Rogers, C., Trick, M., Morieri, G., Mysore, K. S., Wen, J., Oldroyd, G. E. D., Downie, J. A., & Murray, J. D. (2014). The root hair "infectome" of *Medicago truncatula* uncovers changes in cell cycle genes and reveals a requirement for auxin signaling in rhizobial infection. *Plant Cell*, *26*(12), 4680–4701. https://doi.org/10.1105/tpc.114.133496

Brewin, N. J. (2004). Plant cell wall remodelling in the rhizobium-legume symbiosis. *Critical Reviews in Plant Sciences*, *23*(4), 293–316. https://doi.org/10.1080/07352680490480734

Broghammer, A., Krusell, L., Blaise, M., Sauer, J., Sullivan, J. T., Maolanon, N., Vinther, M., Lorentzen, A., Madsen, E. B., Jensen, K. J., Roepstorff, P., Thirup, S., Ronson, C. W., Thygesen, M. B., & Stougaard, J. (2012). Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(34), 13859–13864. https://doi.org/10.1073/pnas.1205171109

Brottier, L., Chaintreuil, C., Simion, P., Scornavacca, C., Rivallan, R., Mournet, P., Moulin, L., Lewis, G. P., Fardoux, J., Brown, S. C., Gomez-Pacheco, M., Bourges, M., Hervouet, C., Gueye, M., Duponnois, R., Ramanankierana, H., Randriambanona, H., Vandrot, H., Zabaleta, M., … Arrighi, J. F. (2018). A phylogenetic framework of the legume genus *Aeschynomene* for comparative genetic analysis of the Nod-dependent and Nod-independent symbioses. *BMC Plant Biology*, *18*(1), 1–15. https://doi.org/10.1186/s12870-018-1567-z

Brown, D. E., Rashotte, A. M., Murphy, A. S., Normanly, J., Tague, B. W., Peer, W. A., Taiz, L., & Muday, G. K. (2001). Flavonoids act as negative regulators of auxin transport *in vivo* in *Arabidopsis*. *Plant Physiology*, *126*(2), 524–535. https://doi.org/10.1104/pp.126.2.524

Busset, N., Gully, D., Teulet, A., Fardoux, J., Camuel, A., Cornu, D., Severac, D., Giraud, E., & Mergaert, P. (2021). The type III effectome of the symbiotic *Bradyrhizobium vignae* strain ORS3257. *Biomolecules*, *11*(11), 1–19. https://doi.org/10.3390/biom11111592

Camuel, A., Teulet, A., Carcagno, M., Haq, F., Pacquit, V., Gully, D., Pervent, M., Chaintreuil, C., Fardoux, J., Horta-araujo, N., Okazaki, S., Ratu, N., & Gueye, F. (2023). Widespread *Bradyrhizobium* distribution of diverse Type III effectors that trigger legume nodulation in the absence of Nod factor. *ISME Journal*, *17*(April), 1416–1429. https://doi.org/10.1038/s41396-023-01458-1

Capoen, W., Oldroyd, G., Goormachtig, S., & Holsters, M. (2010). *Sesbania rostrata*: a case study of natural variation in legume nodulation. *New Phytologist*, *186*(2), 340–345. https://doi.org/10.1111/j.1469-8137.2009.03124.x

Capoen, W., Sun, J., Wysham, D., Otegui, M. S., Venkateshwaran, M., Hirsch, S., Miwa, H., Downie, J. A., Morris, R. J., Ané, J. M., & Oldroyd, G. E. D. (2011). Nuclear membranes control symbiotic calcium signaling of legumes. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(34), 14348–14353. https://doi.org/10.1073/pnas.1107912108

Cárdenas, L., Vidali, L., Domínguez, J., Pérez, H., Sánchez, F., Hepler, P. K., & Quinto, C. (1998). Rearrangement of Actin Microfilaments in Plant Root Hairs Responding to *Rhizobium etli* Nodulation Signals. *Plant Physiology*, *116*(3), 871–877. https://doi.org/10.1104/pp.116.3.871

Cassan, F., & Diaz-Zorita, M. (2016). *Azospirillum* sp. in current agriculture: From thhe laboratory to the field. *Soil Biology & Biochemistry*, *103*, 117–130. https://doi.org/10.1016/j.soilbio.2016.08.020

Centeno-González, N. K., Martínez-Cabrera, H. I., Porras-Múzquiz, H., & Estrada-Ruiz, E. (2021). Late Campanian fossil of a legume fruit supports Mexico as a center of Fabaceae radiation. *Communications Biology*, *4*(1), 1–8. https://doi.org/10.1038/s42003-020-01533-9

Cerri, M. R., Wang, Q., Stolz, P., Folgmann, J., Frances, L., Katzer, K., Li, X., Heckmann, A. B., Wang, T. L., Downie, J. A., Klingl, A., de Carvalho-Niebel, F., Xie, F., & Parniske, M. (2017). The *ERN1* transcription factor gene is a target of the CCaMK/CYCLOPS complex and controls rhizobial infection in *Lotus japonicus*. *New Phytologist*, *215*(1), 323–337. https://doi.org/10.1111/nph.14547

Chaintreuil, C., Giraud, E., Prin, Y., Lorquin, J., Bâ, A., Gillis, M., De Lajudie, P., & Dreyfus, B. (2000). Photosynthetic bradyrhizobia are natural endophytes of the African wild rice *Oryza breviligulata*. *Applied and Environmental Microbiology*, *66*(12), 5437–5447. https://doi.org/10.1128/AEM.66.12.5437-5447.2000

Chaintreuil, Clémence, Arrighi, J. F., Giraud, E., Miché, L., Moulin, L., Dreyfus, B., Munive-Hernández, J. A., Villegas-Hernandez, M. D. C., & Béna, G. (2013). Evolution of symbiosis in the legume genus *Aeschynomene*. *New Phytologist*, *200*(4), 1247–1259. https://doi.org/10.1111/nph.12424

Chaintreuil, Clémence, Rivallan, R., Bertioli, D. J., Klopp, C., Gouzy, J., Courtois, B., Leleux, P., Martin, G., Rami, J. F., Gully, D., Parrinello, H., Séverac, D., Patrel, D., Fardoux, J., Ribière, W., Boursot, M., Cartieaux, F., Czernic, P., Ratet, P., … Arrighi, J. F. (2016). A gene-based map of the Nod factor independent *Aeschynomene evenia* sheds new light on the evolution of nodulation and legume genomes. *DNA Research*, *23*(4), 365–376. https://doi.org/10.1093/dnares/dsw020

Chang, C., Damiani, I., Puppo, A., & Frendo, P. (2009). Redox changes during the legume-rhizobium symbiosis. *Molecular Plant*, *2*(3), 370–377. https://doi.org/10.1093/mp/ssn090

Charpentier, M. (2018). Calcium signals in the plant nucleus: Origin and function. *Journal of Experimental Botany*, *69*(17), 4165–4173. https://doi.org/10.1093/jxb/ery160

Charpentier, M., Sun, J., Martins, T. V., Radhakrishnan, G. V, Findlay, K., Soumpourou, E., Thouin, J., Véry, A., Sanders, D., Morris, R. J., & Oldroyd, G. E. D. (2016). Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science*, *352*(6289), 1102–1105. https://doi.org/10.1126/science.aae0109

Chen, L., Qin, L., Zhou, L., Li, X., Chen, Z., Sun, L., Wang, W., Lin, Z., Zhao, J., Yamaji, N., Ma, J. F., Gu, M., Xu, G., & Liao, H. (2019). A nodule-localized phosphate transporter GmPT7 plays an important role in enhancing symbiotic N 2 fixation and yield in soybean. *New Phytologist*, *221*(4), 2013–2025. https://doi.org/10.1111/nph.15541

Chen, W., Chi, Y., Zhang, J., Bai, B., Ji, X., & Shen, Y. (2022). *MtWRP1*, a Novel Fabacean Specific Gene, Regulates Root Nodulation and Plant Growth in *Medicago truncatula*. *Genes*, *13*(2). https://doi.org/10.3390/genes13020193

Choi, I. S., Cardoso, D., de Queiroz, L. P., de Lima, H. C., Lee, C., Ruhlman, T. A., Jansen, R. K., & Wojciechowski, M. F. (2022). Highly Resolved Papilionoid Legume Phylogeny Based on Plastid Phylogenomics. *Frontiers in Plant Science*, *13*(February), 1–22. https://doi.org/10.3389/fpls.2022.823190

Clerissi, C., Touchon, M., Capela, D., Tang, M., Cruveiller, S., Parker, M. A., Moulin, L., Masson-Boivin, C., &

Rocha, E. P. C. (2018). Parallels between experimental and natural evolution of legume symbionts. *Nature Communications*, *9*(1), 1–13. https://doi.org/10.1038/s41467-018-04778-5

Coba de la Peña, T., Fedorova, E., Pueyo, J. J., & Lucas, M. M. (2018). The symbiosome: Legume and rhizobia co-evolution toward a nitrogen-fixing organelle? *Frontiers in Plant Science*, *8*(January), 1–26. https://doi.org/10.3389/fpls.2017.02229

Combier, J. P., Frugier, F., De Billy, F., Boualem, A., El-Yahyaoui, F., Moreau, S., Vernié, T., Ott, T., Gamas, P., Crespi, M., & Niebel, A. (2006). *MtHAP2-1* is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes and Development*, *20*(22), 3084–3088. https://doi.org/10.1101/gad.402806

Complainville, A., Brocard, L., Roberts, I., Dax, E., Sever, N., Kondorosi, A., Wolf, S., Oparka, K., & Crespi, M. (2003). Nodule Initiation Involves the Creation of a New Symplasmic Field in Specific Root Cells of *Medicago* Species. *The Plant Cell*, *15*(12), 2778–2791. https://doi.org/10.1105/tpc.017020.of

Cordoba, E., Shishkova, S., Vance, C. P., & Hernández, G. (2003). Antisense inhibition of NADH glutamate synthase impairs carbon/nitrogen assimilation in nodules of alfalfa (*Medicago sativa* L.). *Plant Journal*, *33*(6), 1037–1049. https://doi.org/10.1046/j.1365-313X.2003.01686.x

Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: Structural and mechanistic insights. *Nature Reviews Microbiology*, *13*(6), 343–359. https://doi.org/10.1038/nrmicro3456

Cubeñas-Potts, C., & Matunis, M. J. (2013). SUMO: a multifaceted modifier of chromatin structure and function. *Developmental Cell*, *141*(4), 520–529. https://doi.org/10.1016/j.devcel.2012.11.020.SUMO

Cui, H., Tsuda, K., & Parker, J. E. (2015). Effector-triggered immunity: From pathogen perception to robust defense. *Annual Review of Plant Biology*, *66*, 487–511. https://doi.org/10.1146/annurev-arplant-050213-040012

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., Pierre, O., Fardoux, J., Chaintreuil, C., Nguyen, P., Gressent, F., Da Silva, C., Poulain, J., Wincker, P., Rofidal, V., Hem, S., Barrière, Q., Arrighi, J. F., Mergaert, P., & Giraud, E. (2015). Convergent evolution of endosymbiont differentiation in dalbergioid and inverted repeat-lacking clade legumes mediated by nodule-specific cysteine-rich peptides. *Plant Physiology*, *169*(2), 1254–1265. https://doi.org/10.1104/pp.15.00584

Daubech, B., Poinsot, V., Klonowska, A., Capela, D., Chaintreuil, C., Moulin, L., Marchetti, M., & Masson-Boivin, C. (2019). *NoeM*, a new nodulation gene involved in the biosynthesis of nod factors with an open-chain oxidized terminal residue and in the symbiosis with *Mimosa pudica*. *Molecular Plant-Microbe Interactions*, *32*(12), 1635–1648. https://doi.org/10.1094/MPMI-06-19-0168-R

David, M., Daveran, M. L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P., & Kahn, D. (1988). Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. *Cell*, *54*(5), 671–683. https://doi.org/10.1016/S0092-8674(88)80012-6

De Lajudie, P. M., Andrews, M., Ardley, J., Eardly, B., Jumas-Bilak, E., Kuzmanović, N., Lassalle, F., Lindström, K., Mhamdi, R., Martínez-Romero, E., Moulin, L., Mousavi, S. A., Nesme, X., Peix, A., Puławska, J., Steenkamp, E., Stępkowski, T., Tian, C. F., Vinuesa, P., … Young, P. (2019). Minimal standards for the description of new genera and species of rhizobia and agrobacteria. *International Journal of Systematic and Evolutionary Microbiology*, *69*(7), 1852–1863. https://doi.org/10.1099/ijsem.0.003426

De Lajudie, P. M., & Young, J. P. W. (2017). International Committee on Systematics of Prokaryotes Subcommittee for the Taxonomy of *Rhizobium* and *Agrobacterium* Minutes of the meeting, Budapest, 25 August 2016. *International Journal of Systematic and Evolutionary Microbiology*, *67*(7), 2485–2494. https://doi.org/10.1099/ijsem.0.002144

de Moura, G. G. D., Remigi, P., Masson-Boivin, C., & Capela, D. (2020). Experimental evolution of legume symbionts: What have we learnt? *Genes*, *11*(3). https://doi.org/10.3390/genes11030339

Deakin, W. J., & Broughton, W. J. (2009). Symbiotic use of pathogenic strategies: Rhizobial protein secretion systems. *Nature Reviews Microbiology*, *7*(4), 312–320. https://doi.org/10.1038/nrmicro2091

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-gonzález, A., Eldridge, D. J., Bardgett, R. D., Maestre, F. T., Singh, B. K., & Fierer, N. (2018). Bacteria Found in Soil. *Science*, *325*(February), 320–325.

Dénarié, J., Debellé, F., & Promé, J. C. (1996). Rhizobium lipo-chitooligosaccharide nodulation factors: Signaling molecules mediating recognition and morphogenesis. *Annual Review of Biochemistry*, *65*, 503–535. https://doi.org/10.1146/annurev.bi.65.070196.002443

Deng, J., Zhu, F., Liu, J., Zhao, Y., Wen, J., Wang, T., & Dong, J. (2019). Transcription factor bHLH2 Represses *CYSTEINE PROTEASE77* to negatively regulate nodule senescence. *Plant Physiology*, *181*(4), 1683–1703. https://doi.org/10.1104/pp.19.00574

Denison, R. F., & Kiers, E. T. (2011). Life histories of symbiotic rhizobia and mycorrhizal fungi. *Current Biology*, *21*(18), R775–R785. https://doi.org/10.1016/j.cub.2011.06.018

Desbrosses, G. J., & Stougaard, J. (2011). Root nodulation: A paradigm for how plant-microbe symbiosis influences host developmental pathways. *Cell Host and Microbe*, *10*(4), 348–358. https://doi.org/10.1016/j.chom.2011.09.005

Domonkos, A., Horvath, B., Marsh, J. F., Halasz, G., Ayaydin, F., Oldroyd, G. E. D., & Kalo, P. (2013). The identification of novel loci required for appropriate nodule development in *Medicago truncatula*. *BMC Plant Biology*, *13*(1), 1. https://doi.org/10.1186/1471-2229-13-157

Dong, W., & Song, Y. (2020). The significance of flavonoids in the process of biological nitrogen fixation. *International Journal of Molecular Sciences*, *21*(16), 1–18. https://doi.org/10.3390/ijms21165926

Downie, J. A. (2005). Legume Haemoglobins: symbiotic nitrogen fixation needs bloody nodules. *Current Biology*, *15*(6), 196–198. https://doi.org/10.1016/j.cub.2005.03.006

Downie, J. A. (2010). The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. *FEMS Microbiology Reviews*, *34*(2), 150–170. https://doi.org/10.1111/j.1574-6976.2009.00205.x

Doyle, J. J. (1998). Phylogenetic perspectives on nodulation: Evolving views of plants and symbiotic bacteria. *Trends in Plant Science*, *3*(12), 473–478. https://doi.org/10.1016/S1360-1385(98)01340-5

Doyle, J. J., & Luckow, M. A. (2003). The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiology*, *131*(3), 900–910. https://doi.org/10.1104/pp.102.018150

Eady, R. R. (1996). Structure-function relationships of alternative nitrogenases. *Chemical Reviews*, *96*(7), 3013–3030. https://doi.org/10.1021/cr950057h

Erisman, J. W., Galloway, J. N., Seitzinger, S., Bleeker, A., Dise, N. B., Roxana Petrescu, A. M., Leach, A. M., & de Vries, W. (2013). Consequences of human modification of the global nitrogen cycle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621). https://doi.org/10.1098/rstb.2013.0116

Erisman, J. W., Sutton, M. A., Galloway, J., Klimont, Z., & Winiwarter, W. (2008). How a century of ammonia synthesis changed the world. *Nature Geoscience*, *1*(10), 636–639. https://doi.org/10.1038/ngeo325

Esseling, J. J., Lhuissier, F. G. P., & Emons, A. M. C. (2003). Nod factor-induced root hair curling: Continuous polar growth towards the point of nod factor application. *Plant Physiology*, *132*(4), 1982–1988. https://doi.org/10.1104/pp.103.021634

Evans, W. R., Fleischman, D., Calvert, H. E., Pyati, P. V, Alter, G. M., & Subba Rao, N. S. (1990). Bacteriochlorophyll and photosynthetic reaction centers in *Rhizobium* strain BTAi1. *Applied and Environmental Microbiology*, *56*(11), 3445–3449. https://doi.org/10.1128/aem.56.11.3445-3449.1990

Fabre, S., Gully, D., Poitout, A., Patrel, D., Arrighi, J. F., Giraud, E., Czernic, P., & Cartieaux, F. (2015). Nod factor-independent nodulation in *Aeschynomene evenia* required the common plant-microbe symbiotic toolkit. *Plant Physiology*, *169*(4), 2654–2664. https://doi.org/10.1104/pp.15.01134

Faruque, O. M., Miwa, H., Yasuda, M., Fujii, Y., Kaneko, T., Sato, S., & Okazaki, S. (2015). Identification of *Bradyrhizobium elkanii* genes involved in incompatibility with soybean plants carrying the *Rj4* allele. *Applied and Environmental Microbiology*, *81*(19), 6710–6717. https://doi.org/10.1128/AEM.01942-15

Feng, Y., Wu, P., Liu, C., Peng, L., Wang, T., Wang, C., Tan, Q., Li, B., Ou, Y., Zhu, H., Yuan, S., Huang, R., Stacey, G., Zhang, Z., & Cao, Y. (2021). Suppression of LjBAK1-mediated immunity by SymRK promotes rhizobial infection in *Lotus japonicus*. *Molecular Plant*, *14*(11), 1935–1950. https://doi.org/10.1016/j.molp.2021.07.016

Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M. H., Lin, Y. H., Reid, D. E., & Gresshoff, P. M. (2010).

Molecular analysis of legume nodule development and autoregulation. *Journal of Integrative Plant Biology*, *52*(1), 61–76. https://doi.org/10.1111/j.1744-7909.2010.00899.x

Ferguson, B. J., & Mathesius, U. (2014). Phytohormone Regulation of Legume-Rhizobia Interactions. *Journal of Chemical Ecology*, *40*(7), 770–790. https://doi.org/10.1007/s10886-014-0472-7

Ferguson, B. J., Mens, C., Hastwell, A. H., Zhang, M., Su, H., Jones, C. H., Chu, X., & Gresshoff, P. M. (2019). Legume nodulation: The host controls the party. *Plant Cell and Environment*, *42*(1), 41–51. https://doi.org/10.1111/pce.13348

Ferraioli, S., Täte, R., Rogato, A., Chiurazzi, M., & Patriarca, E. J. (2004). Development of ectopic roots from abortive nodule primordia. *Molecular Plant-Microbe Interactions*, *17*(10), 1043–1050. https://doi.org/10.1094/MPMI.2004.17.10.1043

Figaj, D., Ambroziak, P., Przepiora, T., & Skorko-Glonek, J. (2019). The role of proteases in the virulence of plant pathogenic bacteria. *International Journal of Molecular Sciences*, *20*(3). https://doi.org/10.3390/ijms20030672

Flemetakis, E., Efrose, R. C., Ott, T., Stedel, C., Aivalakis, G., Udvardi, M. K., & Katinakis, P. (2006). Spatial and temporal organization of sucrose metabolism in *Lotus japonicus* nitrogen-fixing nodules suggests a role for the elusive alkaline/neutral invertase. *Plant Molecular Biology*, *62*(1–2), 53–69. https://doi.org/10.1007/s11103-006-9003-4

Fonouni-Farde, C., Tan, S., Baudin, M., Brault, M., Wen, J., Mysore, K. S., Niebel, A., Frugier, F., & Diet, A. (2016). DELLA-mediated gibberellin signalling regulates Nod factor signalling and rhizobial infection. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms12636

Fournier, J., Timmers, A. C. J., Sieberer, B. J., Jauneau, A., Chabaud, M., & Barker, D. G. (2008). Mechanism of infection thread elongation in root hairs of *Medicago truncatula* and dynamic interplay with associated rhizobial colonization. *Plant Physiology*, *148*(4), 1985–1995. https://doi.org/10.1104/pp.108.125674

Fowler, D., Coyle, M., Skiba, U., Sutton, M. A., Cape, J. N., Reis, S., Sheppard, L. J., Jenkins, A., Grizzetti, B., Galloway, J. N., Vitousek, P., Leach, A., Bouwman, A. F., Butterbach-Bahl, K., Dentener, F., Stevenson, D., Amann, M., & Voss, M. (2013). The global nitrogen cycle in the Twentyfirst century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621). https://doi.org/10.1098/rstb.2013.0164

Foyer, C. H., Lam, H. M., Nguyen, H. T., Siddique, K. H. M., Varshney, R. K., Colmer, T. D., Cowling, W., Bramley, H., Mori, T. A., Hodgson, J. M., Cooper, J. W., Miller, A. J., Kunert, K., Vorster, J., Cullis, C., Ozga, J. A., Wahlqvist, M. L., Liang, Y., Shou, H., … Considine, M. J. (2016). Neglecting legumes has compromised human health and sustainable food production. *Nature Plants*, *2*(8), 1–10. https://doi.org/10.1038/NPLANTS.2016.112

Frona, D., Janos, S., & Harangi-Rakos, M. (2019). The challenge of feeding the world. *MDPI Journal for Sustainability*, *2017*, 3–4.

Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P., & Sutton, M. A. (2008). Transformation of the Nitrogen Cycle : *Science*, *320*(May), 889–892.

Gao, F., Yang, J., Zhai, N., Zhang, C., & Ren, X. (2023). NCR343 is required to maintain the viability of differentiated bacteroids in nodule cells in *Medicago truncatula*. *New Phytologist*. https://doi.org/10.1111/nph.19180

Garrido-Sanz, D., Redondo-Nieto, M., Mongiardini, E., Blanco-Romero, E., Durán, D., Quelas, J. I., Martin, M., Rivilla, R., Lodeiro, A. R., & Althabegoiti, M. J. (2019). Phylogenomic analyses of *Bradyrhizobium* reveal uneven distribution of the lateral and subpolar flagellar systems, which extends to rhizobiales. *Microorganisms*, *7*(2). https://doi.org/10.3390/microorganisms7020050

Gaudioso-Pedraza, R., Beck, M., Frances, L., Kirk, P., Ripodas, C., Niebel, A., Oldroyd, G. E. D., Benitez-Alfonso, Y., & de Carvalho-Niebel, F. (2018). Callose-Regulated Symplastic Communication Coordinates Symbiotic Root Nodule Development. *Current Biology*, *28*(22), 3562-3577.e6. https://doi.org/10.1016/j.cub.2018.09.031

Geurts, R., Xiao, T. T., & Reinhold-Hurek, B. (2016). What Does It Take to Evolve A Nitrogen-Fixing Endosymbiosis? *Trends in Plant Science*, *21*(3), 199–208. https://doi.org/10.1016/j.tplants.2016.01.012

Gil-Díez, P., Tejada-Jiménez, M., León-Mediavilla, J., Wen, J., Mysore, K. S., Imperial, J., & González-Guerrero, M. (2019). MtMOT1.2 is responsible for molybdate supply to *Medicago truncatula* nodules. *Plant Cell and Environment*, *42*(1), 310–320. https://doi.org/10.1111/pce.13388

Giraud, E., Fardoux, J., Fourrier, N., Hannibal, L., Genty, B., Bouyer, P., Dreyfus, B., & Verméglio, A. (2002). Bacteriophytochrome controls photosystem synthesis in anoxygenic bacteria. *Nature*, *417*(May), 202–205. https://doi.org/https://doi.org/10.1038/417202a

Giraud, E., & Fleischman, D. (2004). Nitrogen-fixing symbiosis between photosynthetic bacteria and legumes. *Photosynthesis Research*, *82*(2), 115–130. https://doi.org/10.1007/s11120-004-1768-1

Giraud, E., Hannibal, L., Fardoux, J., Verméglio, A., & Dreyfus, B. (2000). Effect of Bradyrhizobium photosynthesis on stem nodulation of *Aeschynomene sensitiva*. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(26), 14795–14800. https://doi.org/10.1073/pnas.250484097

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J. C., Jaubert, M., Simon, D., Cartieaux, F., Prin, Y., Bena, G., Hannibal, L., Fardoux, J., Kojadinovic, M., Vuillet, L., Lajus, A., Cruveiller, S., Rouy, Z., Mangenot, S., … Sadowsky, M. (2007). Legumes symbioses: Absence of *nod* genes in photosynthetic bradyrhizobia. *Science*, *316*(5829), 1307–1312. https://doi.org/10.1126/science.1139548

Gleason, C., Chaudhuri, S., Yang, T., Muñoz, A., Poovaiah, B. W., & Oldroyd, G. E. D. (2006). Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature*, *441*(7097), 1149–1152. https://doi.org/10.1038/nature04812

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., & Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *Science*, *327*(February), 812–818. https://doi.org/https://doi.org/10.1126/science.1185383

Godiard, L., Lepage, A., Moreau, S., Laporte, D., Verdenaud, M., Timmers, T., & Gamas, P. (2011). MtbHLH1, a bHLH transcription factor involved in *Medicago truncatula* nodule vascular patterning and nodule to plant metabolic exchanges. *New Phytologist*, *191*(2), 391–404. https://doi.org/10.1111/j.1469-8137.2011.03718.x

González-Guerrero, M., Matthiadis, A., Sáez, Á., & Long, T. A. (2014). Fixating on metals: New insights into the role of metals in nodulation and symbiotic nitrogen fixation. *Frontiers in Plant Science*, *5*(FEB), 1–6. https://doi.org/10.3389/fpls.2014.00045

Gonzalez-Rizzo, S., Crespi, M., & Frugier, F. (2006). The *Medicago truncatula* CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with *Sinorhizobium meliloti*. *Plant Cell*, *18*(10), 2680–2693. https://doi.org/10.1105/tpc.106.043778

Goormachtig, S., Capoen, W., James, E. K., & Holsters, M. (2004). Switch from intracellular to intercellular invasion during water stress-tolerant legume nodulation. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(16), 6303–6308. https://doi.org/10.1073/pnas.0401540101

Gourion, B., Berrabah, F., Ratet, P., & Stacey, G. (2015). Rhizobium-legume symbioses: The crucial role of plant immunity. *Trends in Plant Science*, *20*(3), 186–194. https://doi.org/10.1016/j.tplants.2014.11.008

Griesmann, M., Chang, Y., Liu, X., Song, Y., Haberer, G., Crook, M. B., Billault-Penneteau, B., Lauressergues, D., Keller, J., Imanishi, L., Roswanjaya, Y. P., Kohlen, W., Pujic, P., Battenberg, K., Alloisio, N., Liang, Y., Hilhorst, H., Salgado, M. G., Hocher, V., … Cheng, S. (2018). Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science*, *361*(6398). https://doi.org/10.1126/science.aat1743

Groth, M., Takeda, N., Perry, J., Uchid, H., Dräxl, S., Brachmann, A., Sato, S., Tabata, S., Kawaguchi, M., Wang, T. L., & Parniskea, M. (2010). *NENA*, a *Lotus japonicus* homolog of *Sec13*, is required for rhizodermal infection by arbuscular mycorrhiza fungi and rhizobia but dispensable for cortical endosymbiotic development. *Plant Cell*, *22*(7), 2509–2526. https://doi.org/10.1105/tpc.109.069807

Grunewald, W., De Smet, I., Lewis, D. R., Löfke, C., Jansen, L., Goeminne, G., Vanden Bossche, R., Karimi, M., De Rybel, B., Vanholme, B., Teichmann, T., Boerjan, W., Van Montagu, M. C. E., Gheysen, G., Muday, G. K., Friml, J., & Beeckman, T. (2012). Transcription factor WRKY23 assists auxin distribution patterns during *Arabidopsis* root development through local control on flavonol biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(5), 1554–1559. https://doi.org/10.1073/pnas.1121134109

Guilfoyle, T., Hagen, G., Ulmasov, T., & Murfett, J. (1998). How does auxin turn on genes? *Plant Physiology*, *118*(2), 341–347. https://doi.org/10.1104/pp.118.2.341

Guinel, F. C. (2009). Getting around the legume nodule: I. The structure of the peripheral zone in four nodule types. *Botany*, *87*(12), 1117–1138. https://doi.org/10.1139/B09-074

Gully, D., Czernic, P., Cruveiller, S., Mahé, F., Longin, C., Vallenet, D., François, P., Nidelet, S., Rialle, S., Giraud, E., Arrighi, J. F., DasGupta, M., & Cartieaux, F. (2018). Transcriptome Profiles of Nod Factor-independent Symbiosis in the Tropical Legume *Aeschynomene evenia*. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-29301-0

Hakoyama, T., Niimi, K., Watanabe, H., Tabata, R., Matsubara, J., Sato, S., Nakamura, Y., Tabata, S., Jichun, L., Matsumoto, T., Tatsumi, K., Nomura, M., Tajima, S., Ishizaka, M., Yano, K., Imaizumi-Anraku, H., Kawaguchi, M., Kouchi, H., & Suganuma, N. (2009). Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. *Nature*, *462*(7272), 514–517. https://doi.org/10.1038/nature08594

Hagerup, O. (1928). En hygrofil Baelgplante (*Aeschynomene aspera* L.) med bakterieknolde paa staengelen. Dan. Bot. ark., 15,1.

Haney, C. H., & Long, S. R. (2010). Plant flotillins are required for infection by nitrogen-fixing bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(1), 478–483. https://doi.org/10.1073/pnas.0910081107

Hardoim, P. R., van Overbeek, L. S., Berg, G., Pirttilä, A. M., Compant, S., Campisano, A., Döring, M., & Sessitsch, A. (2015). The Hidden World within Plants: Ecological and Evolutionary Considerations for Defining Functioning of Microbial Endophytes. *Microbiology and Molecular Biology Reviews*, *79*(3), 293–320. https://doi.org/10.1128/mmbr.00050-14

Hassan, S., & Mathesius, U. (2012). The role of flavonoids in root-rhizosphere signalling: Opportunities and challenges for improving plant-microbe interactions. *Journal of Experimental Botany*, *63*(9), 3429–3444. https://doi.org/10.1093/jxb/err430

Hayashi, Makoto, Imaizumi-Anraku, H., Akao, S., & Kawaguchi, M. (2000). Nodule organogenesis in *Lotus japonicus*. *Journal of Plant Research*, *113*(4), 489–495. https://doi.org/10.1007/pl00013959

Hayashi, Masaki, Saeki, Y., Haga, M., Harada, K., Kouchi, H., & Umehara, Y. (2011). *Rj* (*rj*) genes involved in nitrogen-fixing root nodule formation in soybean. *Breeding Science*, *61*(5), 544–553. https://doi.org/10.1270/jsbbs.61.544

Hayatsu, M., Tago, K., & Saito, M. (2008). Various players in the nitrogen cycle: Diversity and functions of the microorganisms involved in nitrification and denitrification. *Soil Science and Plant Nutrition*, *54*(1), 33–45. https://doi.org/10.1111/j.1747-0765.2007.00195.x

Heidstra, R., Yang, W. C., Yalcin, Y., Peck, S., Emons, A. M., Van Kammen, A., & Bisseling, T. (1997). Ethylene provides positional information on cortical cell division but is not involved in Nod factor-induced root hair tip growth in Rhizobium-legume interaction. *Development*, *124*(9), 1781–1787. https://doi.org/10.1242/dev.124.9.1781

Helene, L. C. F., Delamuta, J. R. M., Ribeiro, R. A., & Hungria, M. (2017). *Bradyrhizobium mercantei* sp. Nov., a nitrogen-fixing symbiont isolated from nodules of *Deguelia costata* (syn. *Lonchocarpus costatus*). *International Journal of Systematic and Evolutionary Microbiology*, *67*(6), 1827–1834. https://doi.org/10.1099/ijsem.0.001870

Herder, J. Den, Vanhee, C., Rycke, R. De, Corich, V., & Holsters, M. (2007). Nod factore perception during infection thread growth fine-tunes nodulation. *Molecular Plant-Microbe Interactions : MPMI*, *20*(2), 129–137. https://doi.org/10.1094 /MPMI -20-2-0129

Herridge, D. F., Peoples, M. B., & Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. *Plant and Soil*, *311*(1–2), 1–18. https://doi.org/10.1007/s11104-008-9668-3

Hocher, V., Alloisio, N., Bogusz, D., & Normand, P. (2011). Early signaling in actinorhizal symbioses. *Plant Signaling and Behavior*, *6*(9), 1377–1379. https://doi.org/10.4161/psb.6.9.16761

Hohnjec, N., Perlick, A. M., Pühler, A., & Küster, H. (2003). The *Medicago truncatula* sucrose synthase gene *MtSucS1* is activated both in the infected region of root nodules and in the cortex of roots colonized by arbuscular mycorrhizal fungi. *Molecular Plant-Microbe Interactions*, *16*(10), 903–915.

https://doi.org/10.1094/MPMI.2003.16.10.903

Hotson, A., Chosed, R., Shu, H., Orth, K., & Mudgett, M. B. (2003). Xanthomonas type III effector XopD targets SUMO-conjugated proteins *in planta*. *Molecular Microbiology*, *50*(2), 377–389. https://doi.org/10.1046/j.1365-2958.2003.03730.x

Hou, S., Liu, Z., Shen, H., & Wu, D. (2019). Damage-associated molecular pattern-triggered immunity in plants. *Frontiers in Plant Science*, *10*(May). https://doi.org/10.3389/fpls.2019.00646

Hou, X., Zhou, J., Liu, C., Liu, L., Shen, L., & Yu, H. (2014). Nuclear factor Y-mediated H3K27me3 demethylation of the *SOC1* locus orchestrates flowering responses of *Arabidopsis*. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms5601

Hu, S. P., Li, J. J., Dhar, N., Li, J. P., Chen, J. Y., Jian, W., Dai, X. F., & Yang, X. Y. (2021). Lysin motif (Lysm) proteins: Interlinking manipulation of plant immunity and fungi. *International Journal of Molecular Sciences*, *22*(6), 1–12. https://doi.org/10.3390/ijms22063114

Hu, Z., Wang, R., Zheng, M., Liu, X., Meng, F., Wu, H., Yao, Y., Xin, M., Peng, H., Ni, Z., & Sun, Q. (2018). TaWRKY51 promotes lateral root formation through negative regulation of ethylene biosynthesis in wheat (*Triticum aestivum* L.). *Plant Journal*, *96*(2), 372–388. https://doi.org/10.1111/tpj.14038

Huang, J., Huang, J., Feng, Q., Shi, Y., Wang, F., Zheng, K., Huang, Q., Jiang, J., Luo, S., Xie, Y., Han, D., Lai, J., & Yang, C. (2023). SUMOylation facilitates the assembly of a Nuclear Factor-Y complex to enhance thermotolerance in *Arabidopsis*. *Journal of Integrative Plant Biology*, *65*(3), 692–702. https://doi.org/10.1111/jipb.13396

Ito, J., Fukaki, H., Onoda, M., Li, L., Li, C., Tasaka, M., & Furutani, M. (2016). Auxin-dependent compositional change in Mediator in ARF7- and ARF19-mediated transcription. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(23), 6562–6567. https://doi.org/10.1073/pnas.1600739113

Jaftha, J. B., Strijdom, B. W., & Steyn, P. L. (2002). Characterization of pigmented methylotrophic bacteria which nodulate *Lotononis bainesii*. *Systematic and Applied Microbiology*, *25*(3), 440–449. https://doi.org/10.1078/0723-2020-00124

Jayasinghearachchi, H. S., & Seneviratne, G. (2004). Can mushrooms fix atmospheric nitrogen? *Journal of Biosciences*, *29*(3), 293–296. https://doi.org/10.1007/BF02702611

Jiménez-Guerrero, I., Medina, C., Vinardell, J. M., Ollero, F. J., & López-Baena, F. J. (2022). The Rhizobial Type 3 Secretion System: The Dr. Jekyll and Mr. Hyde in the Rhizobium–Legume Symbiosis. *International Journal of Molecular Sciences*, *23*(19). https://doi.org/10.3390/ijms231911089

Jin, Y., Liu, H., Luo, D., Yu, N., Dong, W., Wang, C., Zhang, X., Dai, H., Yang, J., & Wang, E. (2016). DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms12433

Johnson, P. T. J., Townsend, A. R., Cleveland, C. C., Glibert, P. M., Howarth, R. W., Mckenzie, V. J., Rejmankova, E., & Ward, M. H. (2010). Linking environmental nutrient enrichment and disease emergence in humans and wildlife. *Ecological Applications*, *20*(1), 16–29. https://doi.org/10.1890/08-0633.1

Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*(7117), 323–329.

Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., & Walker, G. C. (2007). How rhizobial symbionts invade plants: The *Sinorhizobium* - *Medicago* model. *Nature Reviews Microbiology*, *5*(8), 619–633. https://doi.org/10.1038/nrmicro1705

Jordan, D. C. (1982). Transfer of *Rhizobium japonicum* Buchanan 1980 to *Bradyrhizobium* gen. nov., a genus of slow-growing, root nodule bacteria from leguminous plants. *International Journal of Systematic Bacteriology*, *32*(1), 136–139. https://doi.org/10.1099/00207713-32-1-136

Kanamori, N., Madsen, L. H., Radutoiu, S., Frantescu, M., Quistgaard, E. M. H., Miwa, H., Downie, J. A., James, E. K., Felle, H. H., Haaning, L. L., Jensen, T. H., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., & Stougaard, J. (2006). A nucleoporin is required for induction of Ca2+ spiking in legume nodule development and essential for rhizobial and fungal symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 359–364. https://doi.org/10.1073/pnas.0508883103

Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpo, S., Tsuruoka, H., Wada, T., Yamada, M.,

& Tabata, S. (2002). Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Research*, *9*(6), 189–197. https://doi.org/10.1093/dnares/9.6.189

Katagiri, F., & Tsuda, K. (2010). Understanding the plant immune system. *Molecular Plant-Microbe Interactions*, *23*(12), 1531–1536. https://doi.org/10.1094/MPMI-04-10-0099

Kawaharada, Y., Kelly, S., Nielsen, M. W., Hjuler, C. T., Gysel, K., Muszyński, A., Carlson, R. W., Thygesen, M. B., Sandal, N., Asmussen, M. H., Vinther, M., Andersen, S. U., Krusell, L., Thirup, S., Jensen, K. J., Ronson, C. W., Blaise, M., Radutoiu, S., & Stougaard, J. (2015). Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature*, *523*(7560), 308–312. https://doi.org/10.1038/nature14611

Ke, D., Fang, Q., Chen, C., Zhu, H., Chen, T., Chang, X., Yuan, S., Kang, H., Ma, L., Hong, Z., & Zhang, Z. (2012). The small GTPase ROP6 interacts with NFR5 and is involved in nodule formation in *Lotus japonicus*. *Plant Physiology*, *159*(1), 131–143. https://doi.org/10.1104/pp.112.197269

Keating, D. H., Willits, M. G., & Long, S. R. (2002). A *Sinorhizobium meliloti* lipopolysaccharide mutant altered in cell surface sulfation. *Journal of Bacteriology*, *184*(23), 6681–6689. https://doi.org/10.1128/JB.184.23.6681-6689.2002

Kim, J. G., Stork, W., & Mudgett, M. B. (2013). *Xanthomonas* type III effector XopD desumoylates tomato transcription factor SlERF4 to suppress ethylene responses and promote pathogen growth. *Cell Host and Microbe*, *13*(2), 143–154. https://doi.org/10.1016/j.chom.2013.01.006

Kim, S., Zeng, W., Bernard, S., Liao, J., Venkateshwaran, M., Ane, J. M., & Jiang, Y. (2019). Ca2+-regulated Ca2+ channels with an RCK gating ring control plant symbiotic associations. *Nature Communications*, *10*(1), 1–12. https://doi.org/10.1038/s41467-019-11698-5

Kiss, E., Oláh, B., Kaló, P., Morales, M., Heckmann, A. B., Borbola, A., Lózsa, A., Kontár, K., Middleton, P., Downie, J. A., Oldroyd, G. E. D., & Endre, G. (2009). LIN, a novel type of U-Box/WD40 Protein, controls early infection by rhizobia in legumes. *Plant Physiology*, *151*(3), 1239–1249. https://doi.org/10.1104/pp.109.143933

Kistner, C., & Parniske, M. (2002). Evolution of signal transduction in intracellular symbiosis. *Trends in Plant Science*, *7*(11), 511–518. https://doi.org/10.1016/S1360-1385(02)02356-7

Kistner, C., Winzer, T., Pitzschke, A., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Webb, K. J., Szczyglowski, K., & Parniske, M. (2005). Seven *Lotus japonicus* genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. *Plant Cell*, *17*(8), 2217–2229. https://doi.org/10.1105/tpc.105.032714

Kitaeva, A. B., Gorshkov, A. P., Kirichek, E. A., Kusakin, P. G., Tsyganova, A. V., & Tsyganov, V. E. (2021). General patterns and species-specific differences in the organization of the tubulin cytoskeleton in indeterminate nodules of three legumes. *Cells*, *10*(5). https://doi.org/10.3390/cells10051012

Koenen, E. J. M., Ojeda, D. I., Bakker, F. T., Wieringa, J. J., Kidner, C., Hardy, O. J., Pennington, R. T., Herendeen, P. S., Bruneau, A., & Hughes, C. E. (2021). The Origin of the Legumes is a Complex Paleopolyploid Phylogenomic Tangle Closely Associated with the Cretaceous-Paleogene (K-Pg) Mass Extinction Event. *Systematic Biology*, *70*(3), 508–526. https://doi.org/10.1093/sysbio/syaa041

Koenen, E. J. M., Ojeda, D. I., Steeves, R., Migliore, J., Bakker, F. T., Wieringa, J. J., Kidner, C., Hardy, O. J., Pennington, R. T., Bruneau, A., & Hughes, C. E. (2020). Large-scale genomic sequence data resolve the deepest divergences in the legume phylogeny and support a near-simultaneous evolutionary origin of all six subfamilies. *New Phytologist*, *225*(3), 1355–1369. https://doi.org/10.1111/nph.16290

Koropatnick, T. A., Engle, J. T., Apicella, M. A., Stabb, E. V, Goldman, W. E., & Mcfall-ngai, M. J. (2004). Microbial Factor-Mediated Development in a Host-Bacterial Mutualism. *Science*, *306*(November), 1186–1189.

Kosslak, R. M., Bookland, R., Barkei, J., Paaren, H. E., & Appelbaum, E. R. (1987). Induction of Bradyrhizobium japonicum common nod genes by isoflavones isolated from *Glycine max*. *Proceedings of the National Academy of Sciences*, *84*(21), 7428–7432. https://doi.org/10.1073/pnas.84.21.7428

Krusell, L., Krause, K., Ott, T., Desbrosses, G., Krämer, U., Sato, S., Nakamura, Y., Tabata, S., James, E. K., Sandal, N., Stougaard, J., Kawaguchi, M., Miyamoto, A., Suganuma, N., & Udvardi, M. K. (2005). The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. *Plant Cell*, *17*(5), 1625–1636. https://doi.org/10.1105/tpc.104.030106

Kulkarni, G., Busset, N., Molinaro, A., Gargani, D., Chaintreuil, C., Silipo, A., Giraud, E., & Newman, D. K. (2015). Specific hopanoid classes differentially affect free-living and symbiotic states of *Bradyrhizobium diazoefficiens*. *MBio*, *6*(5). https://doi.org/10.1128/mBio.01251-15

Lace, B., Su, C., Invernot Perez, D., Rodriguez-Franco, M., Vernié, T., Batzenschlager, M., Egli, S., Liu, C. W., & Ott, T. (2023). RPG acts as a central determinant for infectosome formation and cellular polarization during intracellular rhizobial infections. *ELife*, *12*, 1–33. https://doi.org/10.7554/eLife.80741

Ladha, J. K., Peoples, M. B., Reddy, P. M., Biswas, J. C., Bennett, A., Jat, M. L., & Krupnik, T. J. (2022). Biological nitrogen fixation and prospects for ecological intensification in cereal-based cropping systems. *Field Crops Research*, *283*(April), 108541. https://doi.org/10.1016/j.fcr.2022.108541

Lamouche, F., Bonadé-Bottino, N., Mergaert, P., & Alunni, B. (2019). Symbiotic efficiency of spherical and elongated bacteroids in the *Aeschynomene-Bradyrhizobium* symbiosis. *Frontiers in Plant Science*, *10*(April). https://doi.org/10.3389/fpls.2019.00377

Lascorz, J., Codina-Fabra, J., Reverter, D., & Torres-Rosell, J. (2022). SUMO-SIM interactions: From structure to biological functions. *Seminars in Cell and Developmental Biology*, *132*(November 2021), 193–202. https://doi.org/10.1016/j.semcdb.2021.11.007

Laus, M. C., Logman, T. J., Lamers, G. E., Van Brussel, A. A. N., Carlson, R. W., & Kijne, J. W. (2006). A novel polar surface polysaccharide from *Rhizobium leguminosarum* binds host plant lectin. *Molecular Microbiology*, *59*(6), 1704–1713. https://doi.org/10.1111/j.1365-2958.2006.05057.x

Lavin, M., Pennington, R. T., Klitgaard, B. B., Sprent, J. I., De Lima, H. C., & Gasson, P. E. (2001). The dalbergioid legumes (fabaceae): Delimitation of a pantropical monophyletic clade. *American Journal of Botany*, *88*(3), 503–533. https://doi.org/10.2307/2657116

Lee, S., Reth, A., Meletzus, D., Sevilla, M., & Kennedy, C. (2000). Characterization of a major cluster of *nif*, *fix*, and associated genes in a sugarcane endophyte, *Acetobacter diazotrophicus*. *Journal of Bacteriology*, *182*(24), 7088–7091. https://doi.org/10.1128/JB.182.24.7088-7091.2000

Lefebvre, B., Timmers, T., Mbengue, M., Moreau, S., Hervé, C., Tóth, K., Bittencourt-Silvestre, J., Klaus, D., Deslandes, L., Godiard, L., Murray, J. D., Udvardi, M. K., Raffaele, S., Mongrand, S., Cullimore, J., Gamas, P., Niebel, A., & Ott, T. (2010). A remorin protein interacts with symbiotic receptors and regulates bacterial infection. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(5), 2343–2348. https://doi.org/10.1073/pnas.0913320107

Leitão, N., Dangeville, P., Carter, R., & Charpentier, M. (2019). Nuclear calcium signatures are associated with root development. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-12845-8

León-Mediavilla, J., Senovilla, M., Montiel, J., Gil-Díez, P., Saez, Á., Kryvoruchko, I. S., Reguera, M., Udvardi, M. K., Imperial, J., & González-Guerrero, M. (2018). MtMTP2-Facilitated zinc transport into intracellular compartments is essential for nodule development in *Medicago truncatula*. *Frontiers in Plant Science*, *9*(July), 1–14. https://doi.org/10.3389/fpls.2018.00990

Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J. C., & Dénarié, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature*, *344*(6268), 781–784. https://doi.org/10.1038/344781a0

Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.-P., Ané, J.-M., Lauber, E., Bisseling, T., Dénarié, J., Rosenberg, C., & Debellé, F. (2004). A Putative Ca 2+ and Calmodulin- Dependent Protein Kinase Required for Bacterial and Fungal Symbioses. *Science (New York, N.Y.)*, *303*(5662), 1361–1364. http://www.ncbi.nlm.nih.gov/pubmed/14963335

Lewis, G. ., Schrire, B., Mackinder, B., & Lock, M. (2005). *Legumes of the world* (Royal Botanic Gardens (ed.)).

Li, F., Hou, B., Chen, L., Yao, Z., & Hong, G. (2008). *In vitro* observation of the molecular interaction between NodD and its inducer naringenin as monitored by fluorescence resonance energy transfer. *Acta Biochimica et Biophysica Sinica*, *40*(9), 783–789. https://doi.org/10.1111/j.1745-7270.2008.00462.x

Li, Y., Perez-gil, J., Lois, M., Varejão, N., & Reverter, D. (2023). Multipotent ubiquitin / ubiquitin-like deconjugation activity of the rhizobial effector NopD. *BioRxiv*, 1–30. https://doi.org/https://doi.org/10.1101/2023.09.20.558580

Liang, X., & Zhang, J. (2022). Regulation of plant responses to biotic and abiotic stress by receptor-like

cytoplasmic kinases. *Stress Biology*, *2*(1). https://doi.org/10.1007/s44154-022-00045-2

Libault, M., Farmer, A., Brechenmacher, L., Drnevich, J., Langley, R. J., Bilgin, D. D., Radwan, O., Neece, D. J., Clough, S. J., May, G. D., & Stacey, G. (2010). Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection. *Plant Physiology*, *152*(2), 541–552. https://doi.org/10.1104/pp.109.148379

Libourel, C., Keller, J., Brichet, L., Cazalé, A., Carrère, S., Vernié, T., Couzigou, J., Callot, C., Dufau, I., Cauet, S., Marande, W., Bulach, T., Suin, A., Masson-boivin, C., Remigi, P., Delaux, P., & Capela, D. (2023). Comparative phylotranscriptomics reveals ancestral and derived root nodule symbiosis programmes. *Nature Plants*, *9*(7), 1067–1080. https://doi.org/10.1038/s41477-023-01441-w

Limpens, E., & Bisseling, T. (2014). CYCLOPS: A new vision on rhizobium-induced nodule organogenesis. *Cell Host and Microbe*, *15*(2), 127–129. https://doi.org/10.1016/j.chom.2014.01.015

Ling, J., Wang, H., Wu, P., Li, T., Tang, Y., Naseer, N., Zheng, H., Masson-Boivin, C., Zhong, Z., & Zhu, J. (2016). Plant nodulation inducers enhance horizontal gene transfer of *Azorhizobium caulinodans* symbiosis island. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(48), 13875–13880. https://doi.org/10.1073/pnas.1615121113

Liu, C. W., & Murray, J. D. (2016). The role of flavonoids in nodulation host-range specificity: An update. *Plants*, *5*(3), 4045–4053. https://doi.org/10.3390/plants5030033

Liu, W., Han, X., Zhan, G., Zhao, Z., Feng, Y., & Wu, C. (2015). A novel sucrose-regulatory MADS-box transcription factor GmNMHC5 promotes root development and nodulation in soybean (*Glycine max* [L.] Merr.). *International Journal of Molecular Sciences*, *16*(9), 20657–20673. https://doi.org/10.3390/ijms160920657

Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., & Kahmann, R. (2015). Fungal effectors and plant susceptibility. *Annual Review of Plant Biology*, *66*, 513–545. https://doi.org/10.1146/annurev-arplant-043014-114623

Loh, J., & Stacey, G. (2003). Nodulation gene regulation in *Bradyrhizobium japonicum*: A unique integration of global regulatory circuits. *Applied and Environmental Microbiology*, *69*(1), 10–17. https://doi.org/10.1128/AEM.69.1.10-17.2003

Long, S. R. (1989). Rhizobium-legume nodulation: Life together in the underground. *Cell*, *56*(2), 203–214. https://doi.org/10.1016/0092-8674(89)90893-3

LPWG. (2013). Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon*, *62*(2), 217–248. https://doi.org/10.12705/622.8

LPWG. (2017). A new subfamily classification of the leguminosae based on a taxonomically comprehensive phylogeny. *Taxon*, *66*(1), 44–77. https://doi.org/10.12705/661.3

Madsen, E. B., Antolín-Llovera, M., Grossmann, C., Ye, J., Vieweg, S., Broghammer, A., Krusell, L., Radutoiu, S., Jensen, O. N., Stougaard, J., & Parniske, M. (2011). Autophosphorylation is essential for the *in vivo* function of the *Lotus japonicus* Nod factor receptor 1 and receptor-mediated signalling in cooperation with Nod factor receptor 5. *Plant Journal*, *65*(3), 404–417. https://doi.org/10.1111/j.1365-313X.2010.04431.x

Madsen, L. H., Tirichine, L., Jurkiewicz, A., Sullivan, J. T., Heckmann, A. B., Bek, A. S., Ronson, C. W., James, E. K., & Stougaard, J. (2010). The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nature Communications*, *1*(1). https://doi.org/10.1038/ncomms1009

Mak, A. N.-S., Bradley, P., Bogdanove, A. J., & Stoddard, B. L. (2013). TAL effectors: function, structure, engineering and applications. *Current Opinion in Structural Biology*, *23*(1), 93–99. https://doi.org/10.1016/j.sbi.2012.11.001

Marie, C., Broughton, W. J., & Deakin, W. J. (2001). Rhizobium type III secretion systems: Legume charmers or alarmers? *Current Opinion in Plant Biology*, *4*(4), 336–342. https://doi.org/10.1016/S1369-5266(00)00182-5

Marie, C., Deakin, W. J., Ojanen-Reuhs, T., Diallo, E., Reuhs, B., Broughton, W. J., & Perret, X. (2004). TtsI, a key regulator of *Rhizobium* species NGR234 is required for type III-dependent protein secrection and synthesis of rhamnose-rich polysaccharides. *Molecular Plant-Microbe Interactions*, *17*(9), 958–966.

https://doi.org/10.1094/MPMI.2004.17.9.958

Marín, M., Uversky, V. N., & Ott, T. (2013). Intrinsic disorder in pathogen effectors: Protein flexibility as an evolutionary hallmark in a molecular arms race. *Plant Cell*, *25*(9), 3153–3157. https://doi.org/10.1105/tpc.113.116319

Marino, D., Andrio, E., Danchin, E. G. J., Oger, E., Gucciardo, S., Lambert, A., Puppo, A., & Pauly, N. (2011). A *Medicago truncatula* NADPH oxidase is involved in symbiotic nodule functioning. *New Phytologist*, *189*(2), 580–592. https://doi.org/10.1111/j.1469-8137.2010.03509.x

Martin, F. M., Uroz, S., & Barker, D. G. (2017). Ancestral alliances: Plant mutualistic symbioses with fungi and bacteria. *Science*, *356*(6340). https://doi.org/10.1126/science.aad4501

Martínez-Romero, E. (2009). Controversies in Science Coevolution in Rhizobium-Legume Symbiosis? *DNA and Cell Biology*, *28*(8), 361–370. https://doi.org/10.1089/dna.2009.0863

Masson-Boivin, C., Giraud, E., Perret, X., & Batut, J. (2009). Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends in Microbiology*, *17*(10), 458–466. https://doi.org/10.1016/j.tim.2009.07.004

Mathesius, U. (2018). Flavonoid functions in plants and their interactions with other organisms. *Plants*, *7*(2), 7–9. https://doi.org/10.3390/plants7020030

Mathesius, U., Schlaman, H. R. M., Spaink, H. P., Sautter, C., Rolfe, B. G., & Djordjevic, M. A. (1998). Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivatives of chitin oligosaccharides. *Plant Journal*, *14*(1), 23–34. https://doi.org/10.1046/j.1365-313X.1998.00090.x

Mbengue, M. D., Hervé, C., & Debellé, F. (2020). Nod factor signaling in symbiotic nodulation. *Advances in Botanical Research*, *94*(20), 1–39. https://doi.org/10.1016/bs.abr.2019.10.002 ©

McAdam, E. L., Reid, J. B., & Foo, E. (2018). Gibberellins promote nodule organogenesis but inhibit the infection stages of nodulation. *Journal of Experimental Botany*, *69*(8), 2117–2130. https://doi.org/10.1093/jxb/ery046

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A. E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., & Kondorosi, E. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(13), 5230–5235. https://doi.org/10.1073/pnas.0600912103

Mergaert, P., Van Montagu, M., & Holsters, M. (1997). Molecular mechanisms of Nod factor diversity. *Molecular Microbiology*, *25*(5), 811–817. https://doi.org/10.1111/j.1365-2958.1997.mmi526.x

Messinese, E., Mun, J. H., Li, H. Y., Jayaraman, D., Rougé, P., Barre, A., Lougnon, G., Schornack, S., Bono, J. J., Cook, D. R., & Ané, J. M. (2007). A novel nuclear protein interacts with the symbiotic DMI3 calcium- and calmodulin-dependent protein kinase of *Medicago truncatula*. *Molecular Plant-Microbe Interactions*, *20*(8), 912–921. https://doi.org/10.1094/MPMI-20-8-0912

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods*, *19*(6), 679–682. https://doi.org/10.1038/s41592-022-01488-1

Mitra, R. M., Gleason, C. A., Edwards, A., Hadfield, J., Downie, J. A., Oldroyd, G. E. D., & Long, S. R. (2004). A Ca2+/calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(13), 4701–4705. https://doi.org/10.1073/pnas.0400595101

Montiel, J., Downie, J. A., Farkas, A., Bihari, P., Herczeg, R., Bálint, B., Mergaert, P., Kereszt, A., & Kondorosi, É. (2017). Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(19), 5041–5046. https://doi.org/10.1073/pnas.1704217114

Montiel, J., Nava, N., Cárdenas, L., Sánchez-López, R., Arthikala, M. K., Santana, O., Sánchez, F., & Quinto, C. (2012). A *Phaseolus vulgaris* NADPH Oxidase Gene is Required for Root Infection by Rhizobia. *Plant and Cell Physiology*, *53*(10), 1751–1767. https://doi.org/10.1093/pcp/pcs120

Morieri, G., Martinez, E. A., Jarynowski, A., Driguez, H., Morris, R., Oldroyd, G. E. D., & Downie, J. A. (2013).

Host-specific Nod-factors associated with *Medicago truncatula* nodule infection differentially induce calcium influx and calcium spiking in root hairs. *New Phytologist*, *200*(3), 656–662. https://doi.org/10.1111/nph.12475

Mornico, D., Miché, L., Béna, G., Nouwen, N., Verméglio, A., Vallenet, D., Smith, A. A. T., Giraud, E., Médigue, C., & Moulin, L. (2012). Comparative genomics of *Aeschynomene* symbionts: Insights into the ecological lifestyle of nod-independent photosynthetic bradyrhizobia. *Genes*, *3*(1), 35–61. https://doi.org/10.3390/genes3010035

Moura, E. G., Carvalho, C. S., Bucher, C. P. C., Souza, J. L. B., Aguiar, A. C. F., Ferraz, A. S. L., Bucher, C. A., & Coelho, K. P. (2020). Diversity of rhizobia and importance of their interactions with legume trees for feasibility and sustainability of the tropical agrosystems. *Diversity*, *12*(5). https://doi.org/10.3390/D12050206

Mousavi, S. A., & Young, J. P. W. (2022). International Committee on Systematics of Prokaryotes, Subcommittee on the taxonomy of Rhizobia and Agrobacteria, minutes of the annual meeting by videoconference, 5 July 2021, followed by online discussion until 31 December 2021. *International Journal of Systematic and Evolutionary Microbiology*, *72*(6), 1–8. https://doi.org/10.1099/ijsem.0.005453

Müller, J., Wiemken, A., & Boller, T. (2001). Redifferentiation of bacteria isolated from *Lotus japonicus* root nodules colonized by *Rhizobium* sp. NGR234. *Journal of Experimental Botany*, *52*(364), 2181–2186. https://doi.org/10.1093/jexbot/52.364.2181

Mulley, G., White, J. P., Karunakaran, R., Prell, J., Bourdes, A., Bunnewell, S., Hill, L., & Poole, P. S. (2011). Mutation of GOGAT prevents pea bacteroid formation and N2 fixation by globally downregulating transport of organic nitrogen sources. *Molecular Microbiology*, *80*(1), 149–167. https://doi.org/10.1111/j.1365-2958.2011.07565.x

Murray, J. D., Muni, R. R. D., Torres-Jerez, I., Tang, Y., Allen, S., Andriankaja, M., Li, G., Laxmi, A., Cheng, X., Wen, J., Vaughan, D., Schultze, M., Sun, J., Charpentier, M., Oldroyd, G., Tadege, M., Ratet, P., Mysore, K. S., Chen, R., & Udvardi, M. K. (2011). *Vapyrin*, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of *Medicago truncatula*. *Plant Journal*, *65*(2), 244–252. https://doi.org/10.1111/j.1365-313X.2010.04415.x

Mus, F., Crook, M. B., Garcia, K., Costas, A. G., Geddes, B. A., Kouri, E. D., Paramasivan, P., Ryu, M. H., Oldroyd, G. E. D., Poole, P. S., Udvardi, M. K., Voigt, C. A., Ané, J. M., & Peters, J. W. (2016). Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. *Applied and Environmental Microbiology*, *82*(13), 3698–3710. https://doi.org/10.1128/AEM.01055-16

Nakagawa, T., & Imaizumi-Anraku, H. (2015). Rice arbuscular mycorrhiza as a tool to study the molecular mechanisms of fungal symbiosis and a potential target to increase productivity. *Rice*, *8*(1). https://doi.org/10.1186/s12284-015-0067-0

Ndungu, S. M., Messmer, M. M., Ziegler, D., Gamper, H. A., Mészáros, É., Thuita, M., Vanlauwe, B., Frossard, E., & Thonar, C. (2018). Cowpea (*Vigna unguiculata* L. Walp) hosts several widespread bradyrhizobial root nodule symbionts across contrasting agro-ecological production areas in Kenya. *Agriculture, Ecosystems and Environment*, *261*(December 2016), 161–171. https://doi.org/10.1016/j.agee.2017.12.014

Newman, J. D., Diebold, R. J., Schultz, B. W., & Noel, K. D. (1994). Infection of soybean and pea nodules by *Rhizobium* spp. purine auxotrophs in the presence of 5-aminoimidazole-4-carboxamide riboside. *Journal of Bacteriology*, *176*(11), 3286–3294. https://doi.org/10.1128/jb.176.11.3286-3294.1994

Ng, J. L. P., Hassan, S., Truong, T. T., Hocart, C. H., Laffont, C., Frugier, F., & Mathesiusa, U. (2015). Flavonoids and auxin transport inhibitors rescue symbiotic nodulation in the *Medicago truncatula* cytokinin perception mutant *cre1*. *Plant Cell*, *27*(8), 2210–2226. https://doi.org/10.1105/tpc.15.00231

Nguyen, H. P., Miwa, H., Kaneko, T., Sato, S., & Okazaki, S. (2017). Identification of *Bradyrhizobium elkanii* genes involved in incompatibility with *Vigna radiata*. *Genes*, *8*(12). https://doi.org/10.3390/genes8120374

Nguyen, Q. M., Iswanto, A. B. B., Son, G. H., & Kim, S. H. (2021). Recent Advances in Effector-Triggered Immunity in Plants: New Pieces in the Puzzle Create a Different Paradigm. *International Journal of Molecular Sciences*, *22*(9). https://doi.org/10.3390/ijms22094709

Nishida, H., Tanaka, S., Handa, Y., Ito, M., Sakamoto, Y., Matsunaga, S., Betsuyaku, S., Miura, K., Soyano, T., Kawaguchi, M., & Suzaki, T. (2018). A NIN-LIKE PROTEIN mediates nitrate-induced control of root nodule symbiosis in *Lotus japonicus*. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-02831-x

Nissan, G., Manulis-Sasson, S., Weinthal, D., Mor, H., Sessa, G., & Barash, I. (2006). The type III effectors HsvG and HsvB of gall-forming *Pantoea agglomerans* determine host specificity and function as transcriptional activators. *Molecular Microbiology*, *61*(5), 1118–1131. https://doi.org/10.1111/j.1365-2958.2006.05301.x

Noel, K. D., Diebold, R. J., Cava, J. R., & Brink, B. A. (1988). Rhizobial purine and pyrimidine auxotrophs: nutrient supplementation, genetic analysis, and the symbiotiic requirement for de novo purine biosynthesis. *Archives of Microbiology*, *149*, 499–506. https://doi.org/https://doi.org/10.1007/BF00446751

Nouwen, N., Arrighi, J. F., Cartieaux, F., Chaintreuil, C., Gully, D., Klopp, C., & Giraud, E. (2017). The role of rhizobial (NifV) and plant (FEN1) homocitrate synthases in *Aeschynomene*/photosynthetic *Bradyrhizobium* symbiosis. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-00559-0

Oka-Kira, E., & Kawaguchi, M. (2006). Long-distance signaling to control root nodule number. *Current Opinion in Plant Biology*, *9*(5), 496–502. https://doi.org/10.1016/j.pbi.2006.07.012

Okazaki, S., Hattori, Y., & Saeki, K. (2007). The *Mesorhizobium loti purB* gene is involved in infection thread formation and nodule development in *Lotus japonicus*. *Journal of Bacteriology*, *189*(22), 8347–8352. https://doi.org/10.1128/JB.00788-07

Okazaki, S., Kaneko, T., Sato, S., & Saeki, K. (2013). Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 17131–17136. https://doi.org/10.1073/pnas.1302360110

Okazaki, Shin, Noisangiam, R., Okubo, T., Kaneko, T., Oshima, K., Hattori, M., Teamtisong, K., Songwattana, P., Tittabutr, P., Boonkerd, N., Saeki, K., Sato, S., Uchiumi, T., Minamisawa, K., & Teaumroong, N. (2015). Genome analysis of a novel *Bradyrhizobium* sp. doa9 carrying a symbiotic plasmid. *PLoS ONE*, *10*(2), 1–18. https://doi.org/10.1371/journal.pone.0117392

Okazaki, Shin, Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., Gully, D., Arrighi, J. F., Furuta, N., Miwa, H., Yasuda, M., Nouwen, N., Teaumroong, N., & Giraud, E. (2016). Rhizobium-legume symbiosis in the absence of Nod factors: Two possible scenarios with or without the T3SS. *ISME Journal*. https://doi.org/10.1038/ismej.2015.103

Okubo, T., Fukushima, S., Itakura, M., Oshima, K., Longtonglang, A., Teaumroong, N., Mitsui, H., Hattori, M., Hattori, R., Hattori, T., & Minamisawa, K. (2013). Genome analysis suggests that the soil oligotrophic bacterium agromonas oligotrophica (*Bradyrhizobium oligotrophicum*) is a nitrogen-fixing symbiont of *Aeschynomene indica*. *Applied and Environmental Microbiology*, *79*(8), 2542–2551. https://doi.org/10.1128/AEM.00009-13

Okubo, T., Fukushima, S., & Minamisawa, K. (2012). Evolution of *Bradyrhizobium-Aeschynomene* mutualism: Living testimony of the ancient world or highly evolved state? *Plant and Cell Physiology*, *53*(12), 2000–2007. https://doi.org/10.1093/pcp/pcs150

Oldroyd, G. E. D. (2013). Speak, friend, and enter: Signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology*, *11*(4), 252–263. https://doi.org/10.1038/nrmicro2990

Oldroyd, G. E. D., & Dixon, R. (2014). Biotechnological solutions to the nitrogen problem. *Current Opinion in Biotechnology*, *26*(December), 19–24. https://doi.org/10.1016/j.copbio.2013.08.006

Oldroyd, G. E. D., Murray, J. D., Poole, P. S., & Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annual Review of Genetics*, *45*, 119–144. https://doi.org/10.1146/annurev-genet-110410-132549

Ormeño-Orrillo, E., & Martínez-Romero, E. (2019). A genomotaxonomy view of the *Bradyrhizobium* genus. *Frontiers in Microbiology*, *10*(JUN), 1–13. https://doi.org/10.3389/fmicb.2019.01334

Ott, T., Van Dongen, J. T., Günther, C., Krusell, L., Desbrosses, G., Vigeolas, H., Bock, V., Czechowski, T., Geigenberger, P., & Udvardi, M. K. (2005). Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. *Current Biology*, *15*(6), 531–535. https://doi.org/10.1016/j.cub.2005.01.042

Parniske, M. (2008). Arbuscular mycorrhiza: The mother of plant root endosymbioses. *Nature Reviews Microbiology*, *6*(10), 763–775. https://doi.org/10.1038/nrmicro1987

Passaglia, L. M. P. (2017). Bradyrhizobium elkanii nod regulon: Insights through genomic analysis. *Genetics and Molecular Biology*, *40*(3), 703–716. https://doi.org/10.1590/1678-4685-gmb-2016-0228

Penmetsa, R. V., & Cook, D. R. (1997). A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. *Science*, *275*(5299), 527–530. https://doi.org/10.1126/science.275.5299.527

Perez-Quintero, A. L., & Szurek, B. (2019). A Decade Decoded: Spies and Hackers in the History of TAL Effectors Research. *Annual Review of Phytopathology*, *57*, 459–481. https://doi.org/10.1146/annurev-phyto-082718-100026

Perret, X., Staehelin, C., & Broughton, W. J. (2000). Molecular Basis of Symbiotic Promiscuity. *Microbiology and Molecular Biology Reviews*, *64*(1), 180–201. https://doi.org/10.1128/mmbr.64.1.180-201.2000

Perrine-Walker, F., Gherbi, H., Imanishi, L., Hocher, V., Ghodhbane-Gtari, F., Lavenus, J., Meriem Benabdoun, F., Nambiar-Veetil, M., Svistoonoff, S., & Laplaze, L. (2011). Symbiotic Signaling in Actinorhizal Symbioses. *Current Protein & Peptide Science*, *999*(999), 1–9. https://doi.org/10.2174/1389211213488422037

Peters, N. K., Frost, J. W., & Long, S. R. (1986). A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science*, *233*(4767), 977–980. https://doi.org/10.1126/science.3738520

Plet, J., Wasson, A., Ariel, F., Le Signor, C., Baker, D., Mathesius, U., Crespi, M., & Frugier, F. (2011). MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in *Medicago truncatula*. *Plant Journal*, *65*(4), 622–633. https://doi.org/10.1111/j.1365-313X.2010.04447.x

Podlešáková, K., Fardoux, J., Patrel, D., Bonaldi, K., Novák, O., Strnad, M., Giraud, E., Spíchal, L., & Nouwen, N. (2013). Rhizobial synthesized cytokinins contribute to but are not essential for the symbiotic interaction between photosynthetic bradyrhizobia and *Aeschynomene* legumes. *Molecular Plant-Microbe Interactions*, *26*(10), 1232–1238. https://doi.org/10.1094/MPMI-03-13-0076-R

Poole, P., Ramachandran, V., & Terpolilli, J. (2018). Rhizobia: From saprophytes to endosymbionts. *Nature Reviews Microbiology*, *16*(5), 291–303. https://doi.org/10.1038/nrmicro.2017.171

Prayitno, J., Rolfe, B. G., & Mathesius, U. (2006). The ethylene-insensitive sickle mutant of *Medicago truncatula* shows altered auxin transport regulation during nodulation. *Plant Physiology*, *142*(1), 168–180. https://doi.org/10.1104/pp.106.080093

Prell, J., White, J. P., Bourdes, A., Bunnewell, S., Bongaerts, R. J., & Poole, P. S. (2009). Legumes regulate *Rhizobium* bacteroid development and persistence by the supply of branched-chain amino acids. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(30), 12477–12482. https://doi.org/10.1073/pnas.0903653106

Price, N. P. J., RelicA, B., Talmont, F., Lewin, A., Promé, D., Pueppke, S. G., Maillet, F., Dénarié, J., Promé, J. -C, & Broughton, W. J. (1992). Broad-host-range *Rhizobium* species strain NGR234 secretes a family of carbamoylated, and fucosylated, nodulation signals that are O-acetylated or sulphated. *Molecular Microbiology*, *6*(23), 3575–3584. https://doi.org/10.1111/j.1365-2958.1992.tb01793.x

Pruneda, J. N., Durkin, C. H., Geurink, P. P., Ovaa, H., Santhanam, B., Holden, D. W., & Komander, D. (2016). The Molecular Basis for Ubiquitin and Ubiquitin-like Specificities in Bacterial Effector Proteases. *Molecular Cell*, *63*(2), 261–276. https://doi.org/10.1016/j.molcel.2016.06.015

Quilbé, J., Lamy, L., Brottier, L., Leleux, P., Fardoux, J., Rivallan, R., Benichou, T., Guyonnet, R., Becana, M., Villar, I., Garsmeur, O., Hufnagel, B., Delteil, A., Gully, D., Chaintreuil, C., Pervent, M., Cartieaux, F., Bourge, M., Valentin, N., … Arrighi, J. F. (2021). Genetics of nodulation in *Aeschynomene evenia* uncovers mechanisms of the rhizobium–legume symbiosis. *Nature Communications*, *12*(1), 1–14. https://doi.org/10.1038/s41467-021-21094-7

Quilbé, J., Montiel, J., Arrighi, J. F., & Stougaard, J. (2022). Molecular Mechanisms of Intercellular Rhizobial Infection: Novel Findings of an Ancient Process. *Frontiers in Plant Science*, *13*(June). https://doi.org/10.3389/fpls.2022.922982

Quilbé, J., Nouwen, N., Pervent, M., Guyonnet, R., Cullimore, J., Gressent, F., Araújo, N. H., Gully, D., Klopp, C., Giraud, E., & Arrighi, J. F. (2022). A mutant-based analysis of the establishment of Nod-independent

symbiosis in the legume *Aeschynomene evenia*. *Plant Physiology*, *190*(2), 1400–1417. https://doi.org/10.1093/plphys/kiac325

Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Grønlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., & Stougaard, J. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature*, *425*(6958), 585–592. https://doi.org/10.1038/nature02039

Radutoiu, S., Madsen, L. H., Madsen, E. B., Jurkiewicz, A., Fukai, E., Quistgaard, E. M. H., Albrektsen, A. S., James, E. K., Thirup, S., & Stougaard, J. (2007). LysM domains mediate lipochitin-oligosaccharide recognition and *Nfr* genes extend the symbiotic host range. *EMBO Journal*, *26*(17), 3923–3935. https://doi.org/10.1038/sj.emboj.7601826

Rahimlou, S., Bahram, M., & Tedersoo, L. (2021). Phylogenomics reveals the evolution of root nodulating alpha- and beta-Proteobacteria (rhizobia). *Microbiological Research*, *250*(February), 126788. https://doi.org/10.1016/j.micres.2021.126788

Raina, A., Khan, S., Tomlekova, N. B., & Ravelombola, W. (2023). Editorial : Legume breeding in transition : innovation and outlook. *Frontiers in Genetics*, *14*(May), 1–3. https://doi.org/10.3389/fgene.2023.1221551

Rapson, T. D., Gregg, C. M., Allen, R. S., Ju, H. K., Doherty, C. M., Mulet, X., Giddey, S., & Wood, C. C. (2020). Insights into Nitrogenase Bioelectrocatalysis for Green Ammonia Production. *ChemSusChem*, *13*(18), 4856–4865. https://doi.org/10.1002/cssc.202001433

Rasmussen, M. W., Roux, M., Petersen, M., & Mundy, J. (2012). MAP kinase cascades in *Arabidopsis* innate immunity. *Frontiers in Plant Science*, *3*(JUL), 1–6. https://doi.org/10.3389/fpls.2012.00169

Ratu, S. T. N., Hirata, A., Kalaw, C. O., Yasuda, M., Tabuchi, M., & Okazaki, S. (2021). Multiple Domains in the Rhizobial Type III Effector Bel2-5 Determine Symbiotic Efficiency With Soybean. *Frontiers in Plant Science*, *12*(June), 1–12. https://doi.org/10.3389/fpls.2021.689064

Ratu, S. T. N., Teulet, A., Miwa, H., Masuda, S., & Nguyen, H. P. (2021). Rhizobia use a pathogenic - like effector to hijack leguminous nodulation signalling. *Scientific Reports*, 1–15. https://doi.org/10.1038/s41598-021-81598-6

Rey, T., Nars, A., Bonhomme, M., Bottin, A., Huguet, S., Balzergue, S., Jardinaud, M. F., Bono, J. J., Cullimore, J., Dumas, B., Gough, C., & Jacquet, C. (2013). NFP, a LysM protein controlling Nod factor perception, also intervenes in *Medicago truncatula* resistance to pathogens. *New Phytologist*, *198*(3), 875–886. https://doi.org/10.1111/nph.12198

Richter, A., Streubel, J., Blücher, C., Szurek, B., Reschke, M., Grau, J., & Boch, J. (2014). A TAL effector repeat architecture for frameshift binding. *Nature Communications*, *5*, 1–10. https://doi.org/10.1038/ncomms4447

Ried, M. K., Antolín-Llovera, M., & Parniske, M. (2014). Spontaneous symbiotic reprogramming of plant roots triggered by receptor-like kinases. *ELife*, *3*(November), 1–17. https://doi.org/10.7554/eLife.03891

Roche, P., Maillet, F., Plazanet, C., Debellé, F., Ferro, M., Truchet, G., Promé, J. C., & Dénarié, J. (1996). The common *nodABC* genes of *Rhizobium meliloti* are host-range determinants. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(26), 15305–15310. https://doi.org/10.1073/pnas.93.26.15305

Rodríguez-Haas, B., Finney, L., Vogt, S., González-Melendi, P., Imperial, J., & González-Guerrero, M. (2013). Iron distribution through the developmental stages of *Medicago truncatula* nodules. *Metallomics*, *5*(9), 1247–1253. https://doi.org/10.1039/c3mt00060e

Roesch, L. F. W., Fulthorpe, R. R., Riva, A., Casella, G., Hadwin, A. K. M., Kent, A. D., Daroub, S. H., Camargo, F. A. O., Farmerie, W. G., & Triplett, E. W. (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME Journal*, *1*(4), 283–290. https://doi.org/10.1038/ismej.2007.53

Rogato, A., D'Apuzzo, E., Barbulova, A., Omrane, S., Stedel, C., Simon-Rosin, U., Katinakis, P., Flemetakis, M., Udvardi, M., & Chiurazzi, M. (2008). Tissue-specific down-regulation of *LjAMT1;1* compromises nodule function and enhances nodulation in *Lotus japonicus*. *Plant Molecular Biology*, *68*(6), 585–595. https://doi.org/10.1007/s11103-008-9394-5

Rossen, L., Shearman, C. A., Johnston, A. W. B., & Downie, J. A. (1985). The *nodD* gene of *Rhizobium*

leguminosarum is autoregulatory and in the presence of plant exudate induces the *nodA,B,C* genes . *The EMBO Journal*, *4*(13A), 3369–3373. https://doi.org/10.1002/j.1460-2075.1985.tb04092.x

Roux, M., Schwessinger, B., Albrecht, C., Chinchilla, D., Jones, A., Holton, N., Malinovsky, F. G., Tör, M., de Vries, S., & Zipfel, C. (2011). The *Arabidopsis* leucine-rich repeat receptor-like kinases BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and biotrophic pathogens. *Plant Cell*, *23*(6), 2440–2455. https://doi.org/10.1105/tpc.111.084301

Roy, D., & Sadanandom, A. (2021). SUMO mediated regulation of transcription factors as a mechanism for transducing environmental cues into cellular signaling in plants. *Cellular and Molecular Life Sciences*, *78*(6), 2641–2664. https://doi.org/10.1007/s00018-020-03723-4

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., Frugoli, J., Dickstein, R., & Udvardi, M. K. (2020). Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *The Plant Cell*, *32*(1), 15–41. https://doi.org/10.1105/tpc.19.00279

Rübsam, H., Krönauer, C., Abel, N. B., Ji, H., Lironi, D., Hansen, S. B., Nadzieja, M., Kolte, M. V., Abel, D., de Jong, N., Madsen, L. H., Liu, H., Stougaard, J., Radutoiu, S., & Andersen, K. R. (2023). Nanobody-driven signaling reveals the core receptor complex in root nodule symbiosis. *Science*, *379*(6629), 272–277. https://doi.org/10.1126/science.ade9204

Saito, K., Yoshikawa, M., Yano, K., Miwa, H., Uchida, H., Asamizu, E., Sato, S., Tabata, S., Imaizumi-Anraku, H., Umehara, Y., Kouchi, H., Murooka, Y., Szczyglowski, K., Downie, J. A., Parniske, M., Hayashi, M., & Kawaguchi, M. (2007). Nucleoporin85 is required for calcium spiking, fungal and bacterial symbioses, and seed production in *Lotus japonicus*. *Plant Cell*, *19*(2), 610–624. https://doi.org/10.1105/tpc.106.046938

Samac, D., Lohar, D. P., Sharopova, N., Endre, G., Pen, S., Town, C., Silverstein, K. A. T., Vandenbosch, K. A., Biology, P., & Pathology, P. (2006). Transcript Analysis of Early Nodulation Events in. *Plant Physiology*, *140*(January), 221–234. https://doi.org/10.1104/pp.105.070326.1

Samalova, M., Brzobohaty, B., & Moore, I. (2005). pOp6/LhGR: A stringently regulated and highly responsive dexamethasone-inducible gene expression system for tobacco. *Plant Journal*, *41*(6), 919–935. https://doi.org/10.1111/j.1365-313X.2005.02341.x

Samalova, M., Kirchhelle, C., & Moore, I. (2019). Universal Methods for Transgene Induction Using the Dexamethasone-Inducible Transcription Activation System pOp6/LhGR in *Arabidopsis* and Other Plant Species. *Current Protocols in Plant Biology*, *4*(2), e20089. https://doi.org/10.1002/cppb.20089

Santi, C., Bogusz, D., & Franche, C. (2013). Biological nitrogen fixation in non-legume plants. *Annals of Botany*, *111*(5), 743–767. https://doi.org/10.1093/aob/mct048

Sapountzis, P., de Verges, J., Rousk, K., Cilliers, M., Vorster, B. J., & Poulsen, M. (2016). Potential for nitrogen fixation in the fungus-growing termite symbiosis. *Frontiers in Microbiology*, *7*(DEC), 1–5. https://doi.org/10.3389/fmicb.2016.01993

Schena, M., Lloyd, A. M., & Davis, R. W. (1991). A steroid-inducible gene expression system for plant cells. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(23), 10421–10425. https://doi.org/10.1073/pnas.88.23.10421

Schiessl, K., Lilley, J. L. S., Lee, T., Tamvakis, I., Kohlen, W., Bailey, P. C., Thomas, A., Luptak, J., Ramakrishnan, K., Carpenter, M. D., Mysore, K. S., Wen, J., Ahnert, S., Grieneisen, V. A., & Oldroyd, G. E. D. (2019). *NODULE INCEPTION* Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in *Medicago truncatula*. *Current Biology*, *29*(21), 3657-3668.e5. https://doi.org/10.1016/j.cub.2019.09.005

Schlesinger, W. H. (2009). On the fate of anthropogenic nitrogen. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(1), 203–208. https://doi.org/10.1073/pnas.0810193105

Schmeisser, C., Liesegang, H., Krysciak, D., Bakkou, N., Le Quéré, A., Wollherr, A., Heinemeyer, I., Morgenstern, B., Pommerening-Röser, A., Flores, M., Palacios, R., Brenner, S., Gottschalk, G., Schmitz, R. A., Broughton, W. J., Perret, X., Strittmatter, A. W., & Streit, W. R. (2009). *Rhizobium* sp. strain NGR234 possesses a remarkable number of secretion systems. *Applied and Environmental Microbiology*, *75*(12), 4035–4045. https://doi.org/10.1128/AEM.00515-09

Schwessinger, B., & Ronald, P. C. (2012). Plant innate immunity: Perception of conserved microbial signatures. *Annual Review of Plant Biology*, *63*, 451–482. https://doi.org/10.1146/annurev-arplant-042811-105518

Shah, L., Sohail, A., Ahmad, R., Cheng, S., Cao, L., & Wu, W. (2022). The Roles of MADS-Box Genes from Root Growth to Maturity in *Arabidopsis* and Rice. *Agronomy*, *12*(3), 1–20. https://doi.org/10.3390/agronomy12030582

Sharma, V., Bhattacharyya, S., Kumar, R., Kumar, A., Gopalakrishnan, S., Dasgupta, M., & Varshney, R. K. (2020). Molecular Basis of Root Nodule Symbiosis between *Bradyrhizobium* and "Crack-Entry" legume groundnut (*Arachis hypogaea* L.). *Plants*, *9*(2), 276. https://doi.org/10.3390/plants9020276

Silipo, A., Vitiello, G., Gully, D., Sturiale, L., Chaintreuil, C., Fardoux, J., Gargani, D., Lee, H. I., Kulkarni, G., Busset, N., Marchetti, R., Palmigiano, A., Moll, H., Engel, R., Lanzetta, R., Paduano, L., Parrilli, M., Chang, W. S., Holst, O., … Molinaro, A. (2014). Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms6106

Singh, J., & Valdés-López, O. (2023). Discovering the genetic modules controlling root nodule symbiosis under abiotic stresses: salinity as a case study. *New Phytologist*, *237*(4), 1082–1085. https://doi.org/10.1111/nph.18627

Singh, S., Katzer, K., Lambert, J., Cerri, M., & Parniske, M. (2014). CYCLOPS, A DNA-binding transcriptional activator, orchestrates symbiotic root nodule development. *Cell Host and Microbe*, *15*(2), 139–152. https://doi.org/10.1016/j.chom.2014.01.011

Smit, G., Swart, S., Lugtenberg, B. J. J., & Kijne, J. W. (1992). Molecular mechanisms of attachment of *Rhizobium* bacteria to plant roots. *Molecular Microbiology*, *6*(20), 2897–2903. https://doi.org/10.1111/j.1365-2958.1992.tb01748.x

Smit, P., Limpens, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C., & Bisseling, T. (2007). *Medicago* LYK3, an entry receptor in rhizobial nodulation factor signaling. *Plant Physiology*, *145*(1), 183–191. https://doi.org/10.1104/pp.107.100495

Songwattana, P., Chaintreuil, C., Wongdee, J., Teulet, A., Mbaye, M., Piromyou, P., Gully, D., Fardoux, J., Zoumman, A. M. A., Camuel, A., Tittabutr, P., Teaumroong, N., & Giraud, E. (2021). Identification of type III effectors modulating the symbiotic properties of *Bradyrhizobium vignae* strain ORS3257 with various *Vigna* species. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-84205-w

Songwattana, P., Noisangiam, R., Teamtisong, K., Prakamhang, J., Teulet, A., Tittabutr, P., Piromyou, P., Boonkerd, N., Giraud, E., & Teaumroong, N. (2017). Type 3 secretion system (T3SS) of *Bradyrhizobium* sp. DOA9 and its roles in legume symbiosis and rice endophytic association. *Frontiers in Microbiology*, *8*(SEP), 1–12. https://doi.org/10.3389/fmicb.2017.01810

Soyano, T., Kouchi, H., Hirota, A., & Hayashi, M. (2013). NODULE INCEPTION Directly Targets *NF-Y* Subunit Genes to Regulate Essential Processes of Root Nodule Development in *Lotus japonicus*. *PLoS Genetics*, *9*(3). https://doi.org/10.1371/journal.pgen.1003352

Soyano, T., Liu, M., Kawaguchi, M., & Hayashi, M. (2021). Leguminous nodule symbiosis involves recruitment of factors contributing to lateral root development. *Current Opinion in Plant Biology*, *59*, 102000. https://doi.org/10.1016/j.pbi.2020.102000

Spaink, H. P. (2000). Root nodulation and infection factors produced by rhizobial bacteria. *Annual Reviews*, *54*, 257–288.

Spaink, H. P., Wijffelman, C. A., Pees, E., Okker, R. J. H., & Lugtenberg, B. J. J. (1987). Rhizobium nodulation gene *nodD* as a determinant of host specificity. *Nature*, *328*(6128), 337–340. https://doi.org/10.1038/328337a0

Sprent, J. (2009). *Legume Nodulation: A Global Perspective.* Wiley-Blackwell.

Sprent, J. I. (1995). Legume trees and shrubs in the tropics: N2 fixation in perspective. *Soil Biology and Biochemistry*, *27*(4–5), 401–407. https://doi.org/10.1016/0038-0717(95)98610-Z

Sprent, J. I., Ardley, J. K., & James, E. K. (2013). From North to South: A latitudinal look at legume nodulation processes. *South African Journal of Botany*, *89*, 31–41. https://doi.org/10.1016/j.sajb.2013.06.011

Sprent, Janet I., Ardley, J., & James, E. K. (2017). Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytologist*, *215*(1), 40–56. https://doi.org/10.1111/nph.14474

Staehelin, C., & Krishnan, H. B. (2015). Nodulation outer proteins: double-edged swords of symbiotic rhizobia.

Biochemical Journal. https://doi.org/10.1042/BJ20150518

Steenkamp, E. T., Stepkowski, T., Przymusiak, A., Botha, W. J., & Law, I. J. (2008). Cowpea and peanut in southern Africa are nodulated by diverse *Bradyrhizobium* strains harboring nodulation genes that belong to the large pantropical clade common in Africa. *Molecular Phylogenetics and Evolution*, *48*(3), 1131–1144. https://doi.org/10.1016/j.ympev.2008.04.032

Stein, L. Y., & Klotz, M. G. (2016). The nitrogen cycle. *Current Biology*, *26*(3), R94–R98. https://doi.org/10.1016/j.cub.2015.12.021

Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Szczyglowski, K., & Parniske, M. (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature*, *417*(6892), 959–962. https://doi.org/10.1038/nature00841

Strozycki, P. M., Szczurek, A., Lotocka, B., Figlerowicz, M., & Legocki, A. B. (2007). Ferritins and nodulation in *Lupinus luteus*: Iron management in indeterminate type nodules. *Journal of Experimental Botany*, *58*(12), 3145–3153. https://doi.org/10.1093/jxb/erm152

Su, C., Zhang, G., Rodriguez-Franco, M., Hinnenberg, R., Wietschorke, J., Liang, P., Yang, W., Uhler, L., Li, X., & Ott, T. (2023). Transcellular progression of infection threads in *Medicago truncatula* roots is associated with locally confined cell wall modifications. *Current Biology*, *33*(3), 533-542.e5. https://doi.org/10.1016/j.cub.2022.12.051

Sullivan, J. T., & Ronson, C. W. (1998). Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(9), 5145–5149. https://doi.org/10.1073/pnas.95.9.5145

Suzaki, T., Takeda, N., Nishida, H., Hoshino, M., Ito, M., Misawa, F., Handa, Y., Miura, K., & Kawaguchi, M. (2019). Lack of symbiont accommodation controls intracellular symbiont accommodation in root nodule and arbuscular mycorrhizal symbiosis in *Lotus Japonicus*. *PLoS Genetics*, *15*(1), 1–25. https://doi.org/10.1371/journal.pgen.1007865

Sy, A., Giraud, E., Jourand, P., Garcia, N., Willems, A., De Lajudie, P., Prin, Y., Neyra, M., Gillis, M., Boivin-Masson, C., & Dreyfus, B. (2001). Methylotrophic *Methylobacterium* bacteria nodulate and fix nitrogen in symbiosis with legumes. *Journal of Bacteriology*, *183*(1), 214–220. https://doi.org/10.1128/JB.183.1.214-220.2001

Takeda, N., Tsuzuki, S., Suzaki, T., Parniske, M., & Kawaguchi, M. (2013). *CERBERUS* and *NSP1* of *Lotus japonicus* are common symbiosis genes that modulate arbuscular mycorrhiza development. *Plant and Cell Physiology*, *54*(10), 1711–1723. https://doi.org/10.1093/pcp/pct114

Tang, D., Wang, G., & Zhou, J. M. (2017). Receptor kinases in plant-pathogen interactions: More than pattern recognition. *Plant Cell*, *29*(4), 618–637. https://doi.org/10.1105/tpc.16.00891

Tao, J., Wang, S., Liao, T., & Luo, H. (2021). Evolutionary origin and ecological implication of a unique nif island in free-living *Bradyrhizobium* lineages. *ISME Journal*, *15*(11), 3195–3206. https://doi.org/10.1038/s41396-021-01002-z

Tejada-Jiménez, M., Castro-Rodríguez, R., Kryvoruchko, I., Mercedes Lucas, M., Udvardi, M., Imperial, J., & González-Guerrero, M. (2015). *Medicago truncatula* Natural Resistance-Associated Macrophage Protein1 is required for iron uptake by rhizobia-infected nodule cells. *Plant Physiology*, *168*(1), 258–272. https://doi.org/10.1104/pp.114.254672

Teulet, A., Busset, N., Fardoux, J., Gully, D., Chaintreuil, C., Cartieaux, F., Jauneau, A., Comorge, V., Okazaki, S., Kaneko, T., Gressent, F., Nouwen, N., Arrighi, J. F., Koebnik, R., Mergaert, P., Deslandes, L., & Giraud, E. (2019). The rhizobial type III effector ErnA confers the ability to form nodules in legumes. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(43), 21758–21768. https://doi.org/10.1073/pnas.1904456116

Teulet, A., Gully, D., Rouy, Z., Camuel, A., Koebnik, R., Giraud, E., & Lassalle, F. (2020). Phylogenetic distribution and evolutionary dynamics of *nod* and *T3SS* genes in the genus *Bradyrhizobium*. *Microbial Genomics*. https://doi.org/10.1099/mgen.0.000407

Teulet, A., Quan, C., Evangelisti, E., Wanke, A., Yang, W., & Schornack, S. (2023). A pathogen effector FOLD diversified in symbiotic fungi. *New Phytologist*, 1127–1139. https://doi.org/10.1111/nph.18996

Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J. G., Jackson, R. B., Arneth, A., Chang, J., Chen, G., Ciais, P., Gerber, S., Ito, A., Huang, Y., Joos, F., Lienert, S., Messina, P., Olin, S., Pan, S., Peng, C., … Zhu, Q. (2018). The global N2O model intercomparison project. *Bulletin of the American Meteorological Society*, *99*(6), 1231–1251. https://doi.org/10.1175/BAMS-D-17-0212.1

Timmers, A. C. J., Auriac, M.-C., & Truchet, G. (1999). Refined analysis of early symbiotic steps of the Rhizobium-Medicago interaction in relationship with microtubular cytoskeleton rearrangements. *Development*, *126*, 3617–3628. file:///C:/PDF/Timmers1999.pdf

Tirichine, L., Imaizumi-Anraku, H., Yoshida, S., Murakami, Y., Madsen, L. H., Miwa, H., Nakagawa, T., Sandal, N., Albrektsen, A. S., Kawaguchi, M., Downie, A., Sato, S., Tabata, S., Kouchi, H., Parniske, M., Kawasaki, S., & Stougaard, J. (2006). Deregulation of a Ca2+/calmodulin-dependent kinase leads to spontaneous nodule development. *Nature*, *441*(7097), 1153–1156. https://doi.org/10.1038/nature04862

Tirichine, L., Sandal, N., Madsen, L. H., Radutoiu, S., Albrektsen, A. S., Sato, S., Asamizu, E., Tabata, S., & Stougaard, J. (2007). A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science*, *315*(5808), 104–107. https://doi.org/10.1126/science.1132397

Tohge, T., De Souza, L. P., & Fernie, A. R. (2017). Current understanding of the pathways of flavonoid biosynthesis in model and crop plants. *Journal of Experimental Botany*, *68*(15), 4013–4028. https://doi.org/10.1093/jxb/erx177

Torsvik, V., & Øvreås, L. (2002). Microbial diversity and function in soil: from genes to ecosystems. *Current Opinion in Microbiology*, *5*, 240–245. https://doi.org/10.1016/s1369-5274(02)00324-7

Townsend, A. R., Howarth, R. W., Bazzaz, F. A., Booth, M. S., Cleveland, C. C., Collinge, S. K., Dobson, A. P., Epstein, P. R., Holland, E. A., Keeney, D. R., Mallin, M. A., Rogers, C. A., Wayne, P., & Wolfe, A. H. (2003). Human Health Effects of a Changing Global Nitrogen Cycle. *Frontiers in Ecology and the Environment*, *1*(5), 240. https://doi.org/10.2307/3868011

Tsyganov, V. E., & Tsyganova, A. V. (2020). Symbiotic regulatory genes controlling nodule development in pisum sativum L. *Plants*, *9*(12), 1–29. https://doi.org/10.3390/plants9121741

Tun-Garrido, C., Bustos, P., González, V., & Brom, S. (2003). Conjugative transfer of p42a from *Rhizobium etli* CFN42, which is required for mobilization of the symbiotic plasmid, is regulated by *quorum sensing*. *Journal of Bacteriology*, *185*(5), 1681–1692. https://doi.org/10.1128/JB.185.5.1681-1692.2003

Udvardi, M., & Poole, P. S. (2013). Transport and metabolism in legume-rhizobia symbioses. *Annual Review of Plant Biology*, *64*, 781–805. https://doi.org/10.1146/annurev-arplant-050312-120235

Vaattovaara, A., Brandt, B., Rajaraman, S., Safronov, O., Veidenberg, A., Luklová, M., Kangasjärvi, J., Löytynoja, A., Hothorn, M., Salojärvi, J., & Wrzaczek, M. (2019). Mechanistic insights into the evolution of DUF26-containing proteins in land plants. *Communications Biology*, *2*(1). https://doi.org/10.1038/s42003-019-0306-9

van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D., Roswanjaya, Y., Wardhani, T. A. K., Kalhor, M. S., Jansen, J., Johan, van den H., Güngör, B., Hartog, M., Hontelez, J., Verver, J., … Geurts, R. (2018). Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(20), E4700–E4709. https://doi.org/10.1073/pnas.1721395115

VanInsberghe, D., Maas, K. R., Cardenas, E., Strachan, C. R., Hallam, S. J., & Mohn, W. W. (2015). Non-symbiotic *Bradyrhizobium* ecotypes dominate North American forest soils. *ISME Journal*, *9*(11), 2435–2441. https://doi.org/10.1038/ismej.2015.54

Veiga, J. P. (2016). Commensalism, Amensalism, and Synnecrosis. *Encyclopedia of Evolutionary Biology*, *October*, 322–328. https://doi.org/10.1016/B978-0-12-800049-6.00189-X

Venkateshwaran, M., Jayaraman, D., Chabaud, M., Genre, A., Balloon, A. J., Maeda, J., Forshey, K., Den Os, D., Kwiecien, N. W., Coon, J. J., Barker, D. G., & Ané, J. M. (2015). A role for the mevalonate pathway in early plant symbiotic signaling. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(31), 9781–9786. https://doi.org/10.1073/pnas.1413762112

Verma, V., Srivastava, A. K., Gough, C., Campanaro, A., Srivastava, M., Morrell, R., Joyce, J., Bailey, M., Zhang,

C., Krysan, P. J., & Sadanandom, A. (2021). SUMO enables substrate selectivity by mitogen-activated protein kinases to regulate immunity in plants. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(10). https://doi.org/10.1073/pnas.2021351118

Vernié, T., Camut, S., Camps, C., Rembliere, C., Carvalho-Niebel, F., Mbengue, M., Timmers, T., Gasciolli, V., Thompson, R., Le Signor, C., Lefebvre, B., Cullimore, J., & Hervé, C. (2016). PUB1 interacts with the receptor kinase DMI2 and negatively regulates rhizobial and arbuscular mycorrhizal symbioses through its ubiquitination activity in *Medicago truncatula*. *Plant Physiology*, *170*(4), 2312–2324. https://doi.org/10.1104/pp.15.01694

Vernié, T., Moreau, S., De Billy, F., Plet, J., Combier, J. P., Rogers, C., Oldroyd, G., Frugier, F., Niebel, A., & Gamas, P. (2008). EFD is an ERF transcription factor involved in the control of nodule number and differentiation in *Medicago truncatula*. *Plant Cell*, *20*(10), 2696–2713. https://doi.org/10.1105/tpc.108.059857

Wais, R. J., Keating, D. H., & Long, S. R. (2002). Structure-function analysis of nod factor-induced root hair calcium spiking in Rhizobium-legume symbiosis. *Plant Physiology*, *129*(1), 211–224. https://doi.org/10.1104/pp.010690

Walker, L., Lagunas, B., & Gifford, M. L. (2020). Determinants of Host Range Specificity in Legume-Rhizobia Symbiosis. *Frontiers in Microbiology*, *11*(November), 1–13. https://doi.org/10.3389/fmicb.2020.585749

Wallner, A. (2021). Genetic Diversity of Type 3 Secretion System in *Burkholderia* s.l. and Links With Plant Host Adaptation. *Frontiers in Microbiology*, *12*(October), 1–13. https://doi.org/10.3389/fmicb.2021.761215

Wang, E. T., Tian, C. F., Chen, W. F., Young, J. P. W., & Chen, W. X. (2019). Ecology and Evolution of Rhizobia. In *Ecology and Evolution of Rhizobia*. https://doi.org/10.1007/978-981-32-9555-1

Wang, H., Chen, W., Xu, Z., Chen, M., & Yu, D. (2023). Functions of WRKYs in plant growth and development. *Trends in Plant Science*, *28*(6), 630–645. https://doi.org/10.1016/j.tplants.2022.12.012

Wang, Q., Huang, Y., Ren, Z., Zhang, X., Ren, J., Su, J., Zhang, C., Tian, J., Yu, Y., Gao, G. F., Li, L., & Kong, Z. (2020). Transfer cells mediate nitrate uptake to control root nodule symbiosis. *Nature Plants*, *6*(7), 800–808. https://doi.org/10.1038/s41477-020-0683-6

Wang, Q., Yang, S., Liu, J., Terecskei, K., Abrahám, E., Gombár, A., Domonkos, A., Szucs, A., Körmöczi, P., Wang, T., Fodor, L., Mao, L., Fei, Z., Kondorosi, E., Kaló, P., Kereszt, A., & Zhu, H. (2017). Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(26), 6854–6859. https://doi.org/10.1073/pnas.1700715114

Wani, S. H., Anand, S., Singh, B., Bohra, A., & Joshi, R. (2021). WRKY transcription factors and plant defense responses: latest discoveries and future prospects. *Plant Cell Reports*, *40*(7), 1071–1085. https://doi.org/10.1007/s00299-021-02691-8

Werner, G. D. A., Cornwell, W. K., Sprent, J. I., Kattge, J., & Kiers, E. T. (2014). A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms. *Nature Communications*, *5*(May), 1–9. https://doi.org/10.1038/ncomms5087

Westhoek, A., Clark, L. J., Culbert, M., Dalchau, N., Griffiths, M., Jorrin, B., Karunakaran, R., Ledermann, R., Tkacz, A., Webb, I., James, E. K., Poole, P. S., & Turnbull, L. A. (2021). Conditional sanctioning in a legume-Rhizobium mutualism. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(19), 1–8. https://doi.org/10.1073/pnas.2025760118

Westhoek, A., Field, E., Rehling, F., Mulley, G., Webb, I., Poole, P. S., & Turnbull, L. A. (2017). Policing the legume-Rhizobium symbiosis: A critical test of partner choice. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-01634-2

Weston, L. A., & Mathesius, U. (2013). Flavonoids: Their Structure, Biosynthesis and Role in the Rhizosphere, Including Allelopathy. *Journal of Chemical Ecology*, *39*(2), 283–297. https://doi.org/10.1007/s10886-013-0248-5

Xiang, Q. W., Bai, J., Cai, J., Huang, Q. Y., Wang, Y., Liang, Y., Zhong, Z., Wagner, C., Xie, Z. P., & Staehelin, C. (2020). NopD of *Bradyrhizobium* sp. XS1150 Possesses SUMO Protease Activity. *Frontiers in Microbiology*, *11*(March), 1–12. https://doi.org/10.3389/fmicb.2020.00386

Xie, F., Murray, J. D., Kim, J., Heckmann, A. B., Edwards, A., Oldroyd, G. E. D., & Downie, J. A. (2012). Legume pectate lyase required for root infection by rhizobia. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(2), 633–638. https://doi.org/10.1073/pnas.1113992109

Xu, L. M., Ge, C., Cui, Z., Li, J., & Fan, H. (1995). *Bradyhizobium liaoningense* sp. nov., Isolated from the Root Nodules of Soybeans MATERIALS AND METHODS. *International Journal of Systematic Bacteriology*, *vi*, 706–711.

Yang, S., Tang, F., Gao, M., Krishnan, H. B., & Zhu, H. (2010). *R* gene-controlled host specificity in the legume-rhizobia symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43), 18735–18740. https://doi.org/10.1073/pnas.1011957107

Yang, Y., Li, L., & Qu, L. J. (2016). Plant Mediator complex and its critical functions in transcription regulation. *Journal of Integrative Plant Biology*, *58*(2), 106–118. https://doi.org/10.1111/jipb.12377

Yano, K., Shibata, S., Chen, W. L., Sato, S., Kaneko, T., Jurkiewicz, A., Sandal, N., Banba, M., Imaizumi-Anraku, H., Kojima, T., Ohtomo, R., Szczyglowski, K., Stougaard, J., Tabata, S., Hayashi, M., Kouchi, H., & Umehara, Y. (2009). CERBERUS, a novel U-box protein containing WD-40 repeats, is required for formation of the infection thread and nodule development in the legume-*Rhizobium* symbiosis. *Plant Journal*, *60*(1), 168–180. https://doi.org/10.1111/j.1365-313X.2009.03943.x

Yano, K., Yoshida, S., Müller, J., Singh, S., Banba, M., Vickers, K., Markmann, K., White, C., Schuller, B., Sato, S., Asamizu, E., Tabata, S., Murooka, Y., Perry, J., Wang, T. L., Kawaguchi, M., Imaizumi-Anraku, H., Hayashi, M., & Parniske, M. (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(51), 20540–20545. https://doi.org/10.1073/pnas.0806858105

Yin, J. W., & Wang, G. (2014). The Mediator complex: A master coordinator of transcription and cell lineage development. *Development (Cambridge)*, *141*(5), 977–987. https://doi.org/10.1242/dev.098392

Yokota, K., Fukai, E., Madsen, L. H., Jurkiewicz, A., Rueda, P., Radutoiu, S., Held, M., Hossain, M. S., Szczyglowski, K., Morieri, G., Oldroyd, G. E. D., Downie, J. A., Nielsen, M. W., Rusek, A. M., Sato, S., Tabata, S., James, E. K., Oyaizu, H., Sandal, N., & Stougaard, J. (2009). Rearrangement of actin cytoskeleton mediates invasion of lotus japonicus roots by *Mesorhizobium loti*. *Plant Cell*, *21*(1), 267–284. https://doi.org/10.1105/tpc.108.063693

Zanetti, M. E., Blanco, F. A., Beker, M. P., Battaglia, M., & Aguilar, O. M. (2010). A C subunit of the plant nuclear factor NF-Y required for rhizobial infection and nodule development affects partner selection in the common bean-*Rhizobium etli* Symbiosis. *Plant Cell*, *22*(12), 4142–4157. https://doi.org/10.1105/tpc.110.079137

Zanetti, M. E., Rípodas, C., & Niebel, A. (2017). Plant NF-Y transcription factors: Key players in plant-microbe interactions, root development and adaptation to stress. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1860*(5), 645–654. https://doi.org/10.1016/j.bbagrm.2016.11.007

Zarrabian, M., Montiel, J., Sandal, N., Ferguson, S., Jin, H., Lin, Y. Y., Klingl, V., Marín, M., James, E. K., Parniske, M., Stougaard, J., & Andersen, S. U. (2022). A Promiscuity Locus Confers *Lotus burttii* Nodulation with Rhizobia from Five Different Genera. *Molecular Plant-Microbe Interactions*, *35*(11), 1006–1017. https://doi.org/10.1094/MPMI-06-22-0124-R

Zhang, R., Wang, Y. H., Jin, J. J., Stull, G. W., Bruneau, A., Cardoso, D., De Queiroz, L. P., Moore, M. J., Zhang, S. D., Chen, S. Y., Wang, J., Li, D. Z., Yi, T. S., & Smith, S. (2020). Exploration of Plastid Phylogenomic Conflict Yields New Insights into the Deep Relationships of Leguminosae. *Systematic Biology*, *69*(4), 613–622. https://doi.org/10.1093/sysbio/syaa013

Zhang, X., Davidson, E. A., Mauzerall, D. L., Searchinger, T. D., Dumas, P., & Shen, Y. (2015). Managing nitrogen for sustainable development. *Nature*, *528*(7580), 51–59. https://doi.org/10.1038/nature15743

Zhao, J., Chen, J., Beillouin, D., Lambers, H., Yang, Y., Smith, P., Zeng, Z., Olesen, J. E., & Zang, H. (2022). Global systematic review with meta-analysis reveals yield advantage of legume-based rotations and its drivers. *Nature Communications*, *13*(1), 1–9. https://doi.org/10.1038/s41467-022-32464-0

Zhou, J. M., & Zhang, Y. (2020). Plant Immunity: Danger Perception and Signaling. *Cell*, *181*(5), 978–989. https://doi.org/10.1016/j.cell.2020.04.028

Zipfel, C. (2009). Early molecular events in PAMP-triggered immunity. *Current Opinion in Plant Biology*, *12*(4), 414–420. https://doi.org/10.1016/j.pbi.2009.06.003

Zipfel, C., & Oldroyd, G. E. D. (2017). Plant signalling in symbiosis and immunity. *Nature*, *543*(7645), 328–336. https://doi.org/10.1038/nature22009

Annexes

Annexe 1 : Articles complémentaires

Article : “Phylogenetic distribution and evolutionary dynamics of nod and T3SS genes in the genus *Bradyrhizobium*”

MICROBIAL GENOMICS

RESEARCH ARTICLE

Teulet *et al.*, *Microbial Genomics* 2020;6
DOI 10.1099/mgen.0.000407

MICROBIOLOGY SOCIETY

OPEN DATA OPEN MICROBIOLOGY

Phylogenetic distribution and evolutionary dynamics of *nod* and T3SS genes in the genus *Bradyrhizobium*

Albin Teulet[1], Djamel Gully[1], Zoe Rouy[2], Alicia Camuel[1], Ralf Koebnik[3], Eric Giraud[1,*,†] and Florent Lassalle[4,5†,‡]

Abstract

Bradyrhizobium are abundant soil bacteria and the major symbiont of legumes. The recent availability of *Bradyrhizobium* genome sequences provides a large source of information for analysis of symbiotic traits. In this study, we investigated the evolutionary dynamics of the nodulation genes (*nod*) and their relationship with the genes encoding type III secretion systems (T3SS) and their effectors among bradyrhizobia. Based on the comparative analysis of 146 *Bradyrhizobium* genome sequences, we identified six different types of T3SS gene clusters. The two predominant cluster types are designated RhcIa and RhcIb and both belong to the RhcI-T3SS family previously described in other rhizobia. They are found in 92/146 strains, most of them also containing *nod* genes. RhcIa and RhcIb gene clusters differ in the genes they carry: while the translocon-encoding gene *nopX* is systematically found in strains containing RhcIb, the *nopE* and *nopH* genes are specifically conserved in strains containing RhcIa, suggesting that these last two genes might functionally substitute *nopX* and play a role related to effector translocation. Phylogenetic analysis suggests that bradyrhizobia simultaneously gained *nod* and RhcI-T3SS gene clusters via horizontal transfer or subsequent vertical inheritance of a symbiotic island containing both. Sequence similarity searches for known Nop effector proteins in bradyrhizobial proteomes revealed the absence of a so-called core effectome, i.e. that no effector is conserved among all *Bradyrhizobium* strains. However, NopM and SUMO proteases were found to be the main effector families, being represented in the majority of the genus. This study indicates that bradyrhizobial T3SSs might play a more significant symbiotic role than previously thought and provides new candidates among T3SS structural proteins and effectors for future functional investigations.

DATA SUMMARY

All the sequence genomes used in this study were downloaded from the National Center for Biotechnology Information (NCBI) RefSeq or GenBank databases. These genomes, which were annotated *de novo* by the MicroScope pipeline, are also publicly available from the MicroScope platform (http://www.genoscope.cns.fr/agc/microscope). Online Supporting Data is available from the Figshare repository at the following http://dx.doi.org/10.6084/m9.figshare.11378772, http://dx.doi.org/10.6084/m9.figshare.12202202 and http://dx.doi.org/10.6084/m9.figshare.12191103.

INTRODUCTION

A recent analysis of microbial communities across the globe indicated that rhizobia of the genus *Bradyrhizobium* are the most ubiquitous and abundant bacteria in soils [1]. They thus appear at the top of the list of the 'most wanted' bacteria to

Received 20 December 2019; Accepted 26 June 2020; Published 12 August 2020
Author affiliations: [1]IRD, Laboratoire des Symbioses Tropicales et Méditerranéennes (LSTM), UMR IRD/SupAgro/INRA/Université de Montpellier/CIRAD, TA-A82/J – Campus de Baillarguet 34398, Montpellier cedex 5, France; [2]LABGeM, Génomique Métabolique, CEA, Genoscope, Institut François Jacob, Université d'Évry, Université Paris-Saclay, CNRS, Evry, France; [3]IRD, CIRAD, Université de Montpellier, IPME, Montpellier, France; [4]Department of Infectious Disease Epidemiology. Imperial College London, St Mary's Hospital Campus, Praed Street, London W2 1NY, UK; [5]Pathogen and Microbes Program, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire CB10 1SA, UK.
***Correspondence:** Eric Giraud, eric.giraud@ird.fr
Keywords: symbiosis; legume; rhizobium; effector; nodulation; type III secretion system.
Abbreviations: Bp, base pair; CDS, protein-coding sequence; DTL, horizontal transfer and loss; ErnA, effector required for nodulation A; ETI, effector-triggered immunity; GBDP, genome-to-genome blast distance phylogeny; HGT, horizontal gene transfer; Mbp, mega base pairs; MIIA, metal ion-inducible autocleavage; ML, maximum likelihood; MSLA, multilocus sequence analysis; NCBI, national center for biotechnology information; NEL, novel E3 ligase; NF, Nod factors; Nop, nodulation outer protein; SUMO, small ubiquitin-like modifier; T3E, type III effector; T3SS, type III secretion system.
‡Present address: Pathogen and Microbes Program, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire CB10 1SA, UK.
†These authors contributed equally to this work
Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Supplementary material is available with the online version of this article.
000407 © 2020 The Authors

This is an open-access article distributed under the terms of the Creative Commons Attribution NonCommercial License.

investigate for a better understanding of the contribution of microbes in the functioning of soil ecosystems. Additionally, these bacteria are the most widely used in agriculture due to their ability to interact symbiotically with legumes of high agronomic importance, including staple crops such as soybean, peanut and cowpea. This symbiosis leads to the formation a new organ, the nodule, generally found on the root system, in which the bacteria fix dinitrogen for the benefit of the plant and receive a carbon source in return. Bradyrhizobia therefore play a very important role in the preservation of the environment by limiting the use of nitrogen fertilizers for the promotion of the growth of legume crops and by contributing to nitrogen enrichment of terrestrial ecosystems.

Not all bradyrhizobia are able to nodulate legumes, as some strains isolated from soils or other environments lack the symbiotic *nod* genes that are necessary for the synthesis of Nod factors (NFs) that govern nodulation [2, 3]. Upon recognition by the plant, this key signal molecule triggers the genetic program leading to nodule formation and its intracellular infection by the bacteria [4]. However, some bradyrhizobia have evolved alternative strategies to interact symbiotically with legumes in the absence of the NF signal. This is the case for some photosynthetic *Bradyrhizobium* strains (including ORS278 and BTAi1), which lack the canonical *nodABC* genes necessary for the synthesis of NFs but are able to induce the formation of nitrogen-fixing nodules on the root and the stem of some legume species of the genus *Aeschynomene* [5]. The bacterial signals necessary for the establishment of this NF-independent symbiosis remain to be discovered.

The *nod* genes, which remain the main determinants conferring the ability to nodulate plants to rhizobia, can be transferred from strain to strain by horizontal gene transfer (HGT) [6, 7]. While these symbiotic genes are typically harboured by a plasmid in other rhizobial genera, these genes are generally found in a chromosomal region of *Bradyrhizobium* strains, called the symbiotic island. In addition, this region generally harbours the *nif* and *fix* genes, which are necessary for the formation and functioning of the nitrogenase complex. This region can be easily recognized, as it is characterized by lower G+C content, different codon usage for the identified protein-coding sequences (CDSs) and the presence of numerous transposases and insertion sequences that strongly promote genomic rearrangements in this region. The high plasticity of the symbiotic island incidentally promotes the selection of symbiotically adaptive variants in bradyrhizobial populations [8]. The presence of genes that code for the type III secretion system (T3SS) has also been frequently reported within this region of *Bradyrhizobium* genomes [9–12].

The T3SS, which was initially found in pathogenic bacteria, is a complex secretory machinery composed of more than 20 conserved proteins that form a syringe-like structure and enable bacteria to deliver type III effector proteins (T3Es) into eukaryotic cells [13]. These T3Es are highly diverse in their structure and mode of action and play a crucial role in bacteria–host interaction by modulating key host cell functions, in most cases related to the suppression of host immunity [14]. Functional T3SSs encoded by 'Rhizobium-conserved' (*rhc*) gene clusters have been reported in several rhizobial species [15, 16]. However, only a few rhizobial T3Es, including those called Nop (nodulation outer proteins), have been characterized and their mode of action generally remains to be determined [16]. The rhizobial T3SS machinery contains most of the core components that were described in the T3SS of pathogenic bacteria, but the genetic organization of the *rhc* cluster and its transcriptional regulation are different from its pathogenic counterpart [17]. Notably, the expression of both *rhc* and *nop* genes is controlled by the regulator TtsI, which is itself under the control of the *nod* gene regulator, NodD [18, 19].

While the virulence of plant pathogenic bacteria largely depends on the functioning of the T3SS [14], the importance of rhizobial T3SS for symbiosis remains unclear. Indeed, the presence of the T3SS among nodulating rhizobia is not a rule, indicating that the T3SS is dispensable for the establishment of the symbiosis. Furthermore, the T3SS can play a neutral, a positive or a negative role, depending on the host plant [16, 20]. On one hand, some of the secreted T3Es may repress the plant immune system, favouring the establishment of the infection and the nodulation. On the other hand, the same effectors may, in other plant contexts, be recognized by receptor proteins [resistant (R) proteins] activating an effector-triggered immunity (ETI), which blocks the

Impact Statement

The nitrogen-fixing symbiosis between legume plants and their bacterial partners (rhizobia) is of immense biological, ecological and agronomic importance, as it is a major contributor to global nitrogen cycling. It has been accepted for a long time that the establishment of the rhizobium–legume symbiosis always relies on the bacterial synthesis of Nod factors, which trigger nodule organogenesis and infection programs. However, recent studies revealed that some rhizobia from the genus *Bradyrhizobium* are able to nodulate some legume species in the absence of Nod factor, relying instead on a type III secretion system (T3SS), a common virulence factor of pathogenic bacteria. In this study, we explored and compared publicly available *Bradyrhizobium* genome sequences to obtain new insights into the significance of the T3SS in the symbiotic adaptation of bradyrhizobia. Our data indicate that the nodulation genes *nod* and T3SS genes have been co-inherited both vertically and horizontally in many instances, and are present together in almost all nodulating *Bradyrhizobium* strains, suggesting an important role of the T3SS in the symbiotic process. Furthermore, we highlighted several genetic determinants that are probably required for the T3SS function and identified newly putative effectors that may contribute to the symbiotic properties of bradyrhizobia.

infection and renders the interaction incompatible [21, 22]. More recently, it has been reported that the T3SS can play a more prominent symbiotic role by directly activating the nodulation in the absence of NFs [23]. This NF-independent, T3SS-dependent symbiotic process was initially discovered between *Bradyrhizobium elkanii* USDA61 and *Glycine max* cv. *Enrei*. This function of the T3SS could be more widespread than originally thought, given that a large and number of diverse non-photosynthetic bradyrhizobia strains were shown to nodulate *Aeschynomene indica* thanks to their T3SS [24]. A recent study conducted on the strain ORS3257 using *A. indica* as host plant revealed that this T3SS-dependent symbiotic process relies on a cocktail of at least five T3Es playing distinct and complementary roles in nodule organogenesis, infection and repression of host immune functions [25]. Among these T3Es, a new one named ErnA (effector required for nodulation A) was shown to act as a key factor to form nodules.

A large number of *Bradyrhizobium* genomes have been sequenced, including those from strains isolated from nodules of several legume species, but also from various other ecological niches. In this study, we took advantage of this wealth of genome sequences to obtain new insights into the significance of the T3SS in the symbiotic adaptation of bradyrhizobia. We examined the phylogenetic distribution of T3SS and *nod* genes as well as some *nop* genes reported to have a symbiotic role to answer to the following questions. What is the distribution of the *nod* and T3SS genes among strains isolated from nodules or other environments? What are the different events of acquisition or loss of the T3SS that occurred during the evolutionary history of bradyrhizobia? Is the acquisition of the T3SS systematically associated with the acquisition of *nod* genes through lateral transfer of the symbiotic island? How congruent is the evolution of the *nod* and T3SS genes? How widespread are the T3Es reported to have a symbiotic role?

METHODS

Phylogenomic analysis of all *Bradyrhizobium* genomes used in this study

A custom set of 155 genomes was gathered covering 28 different *Bradyrhizobium* species, as well as selected genomes from the genera *Nitrobacter* and *Rhodopseudomonas* as outgroups. The genomes listed in Table S1 (available in the online version of this article) were downloaded from the National Center for Biotechnology Information (NCBI) RefSeq or GenBank databases and used as input for the bioinformatic pipeline Pantagruel [26] to build a phylogenomic database. In short, coding sequences (CDSs) and the corresponding protein sequences were extracted from the RefSeq or GenBank annotation and then clustered into homologous gene families using MMSeqs2 [27]. Homologues were aligned with CLUSTAL Omega [28] and reverse-translated into CDS alignment with PAL2NAL [29]. Four hundred and fifty-three single-copy core gene families were selected based on their presence in at least 153 of the 155 genomes, thus allowing for rare losses or incomplete genome sequences. CDS alignments for these families were concatenated, resulting in 378951 aligned nucleotide positions. From this core alignment, a maximum-likelihood (ML) tree was inferred using RAxML (v8.2.4) [30] under the GTRCATX model, and branch lengths were then refined under the GTRGAMMAX model, while branch supports were estimated from 200 rapid bootstrap trees. Bayesian samples of gene trees were inferred from CDS alignments with MrBayes (v 3.2.6) [31] using a Metropolis-coupled Monte Carlo Markov chain (MCMCMC) approach, with two independent sets of four chains (one cold, three heated) running for 2000000 generations and sampled every 500; the presented trees are majority-rule consensus trees, discarding the first 25% of the chain as burn-in. The consensus of Bayesian gene trees for Rhc-T3SS cluster gene families is available as part of the online Supporting Data at the Figshare data repository (http://dx.doi.org/10.6084/m9.figshare.11378772).

Prediction of *nod* and T3SS clusters

All genome sequences used in this study (Table S1) were downloaded using the MicroScope platform (http://www.genoscope.cns.fr/agc/microscope). Genome annotations were performed *de novo* using the MicroScope pipeline [32]. A screen of *nod* and T3SS clusters in each genome was performed by searching the key genes *nodA* and *rhcN*. Reference CDSs were extracted from the *Bradyrhizobium diazoefficiens* USDA110^{T} genome (blr2025 and blr1816, respectively) and searched against the PkGDB genome database using the BLASTN algorithm as implemented on the MicroScope platform with the following parameters: query coverage ≥70%, identity ≥40%. The syntenic organizations of regions surrounding *nodA* and *rhcN* homologues found in bradyrhizobial genomes were then compared with each other using the MaGe genomic visualization system available on the MicroScope platform.

Prediction of known Nops repertoires

The T3Es known to be secreted through the T3SS or to affect the symbiotic interaction were identified from the literature (Table S2) and then searched in all strains using the BLASTP algorithm as implemented on the MicroScope platform with the same similarity constraints as described above. The absence of *nopX* and *nopE*/*nopH* genes in RhcIa- and RhcIb-containing genomes, respectively, were confirmed by TBLASTN search of the NopX proteins from USDA61 strain and NopE/NopH proteins from USDA110^{T}. The search of genes containing known functional domains conserved in some T3Es was performed by interrogating the precomputed InterProScan annotations hosted by the MicroScope platform for each strain genome, using the InterPro identifier of the functional domain as a query.

Phylogenetic analysis

The concatenated NodABC and RhcC2NV amino acid sequences were aligned with CLUSTALW using MEGA X software [33]. Phylogenetic trees were constructed from the resulting alignments using the ML method as implemented in MEGA

X software. Branch support was also estimated using MEGA X by performing bootstrap analyses of 500 replicates. Final editing was carried out using Figtree (http://tree.bio.ed.ac.uk/software/figtree/) and further editing was performed with graphics software. The same procedure was used to construct the SctN phylogenetic tree using the SctN homologues from other T3SS-containing bacteria (*Rhizobiales*, *Enterobacteriales*, *Aeromonadales*, *Pseudomonadales*, *Rhodobacterales*, *Burkholderiales*, *Oceanospirillales*, *Xanthomonadales* and *Chlamydiales*) that are available in the NCBI database.

Prediction of RhcI-T3SS clusters gain and loss events

Evolutionary analyses of the presence/absence of RhcIa- and RhcIb-T3SS clusters were performed using the program Gloome [34]. We used the Mixture model, which allows one to study differences in both gain and loss probabilities. The expectations and probabilities for both gain and loss events were estimated using stochastic mapping and a predicted event was considered as valid if the probability was at least 0.5.

Estimation of individual gene family evolution scenarios and co-evolution

We used gene tree/species tree reconciliation to obtain a finer estimation of the scenarios of events that marked the evolutionary history of individual genes within the RhcI-T3SS clusters and *nod* operons. This approach integrates the information from gene phylogenies to infer events of gene duplication, horizontal transfer and loss (DTL) that occurred over the diversification of the gene family [35]. In short, we used the software GeneRax to estimate ML scenarios as well as to make a Bayesian sample of close-to-optimal scenarios [36]. We applied this procedure to the gene families *nodA*, *nodB*, *nodC*, *rhcC2*, *rhcD*, *rhcN*, *rhcV*, *nopC*, *nopM*, *nopL* and *nopT*, as well as to a sample of control gene families from the bradyrhizobial pangenome; we used the core-genome-based tree of 155 bradyrhizobia and outgroups described above as a reference species tree. Bayesian scenario samples were used to compute co-evolution scores between gene families, using an approach adapted from Lassalle *et al.* [26] and detailed further in File S1. ML and Bayesian samples of scenarios, summaries of the inferred transfer events and of the co-evolution analysis are available as part of online Supporting Data at the Figshare data repository (http://dx.doi.org/10.6084/m9.figshare.12191103). Species tree inference and reconciliation computations were performed on the MRC Cloud Infrastructure for Microbial Bioinformatics (MRC CLIMB) cloud-based computing servers [37].

RESULTS AND DISCUSSION

General features of *Bradyrhizobium* genomes and phylogenomic analysis of the genus

In this study we considered 146 bradyrhizobial genomes and 9 genomes from related genera that were available in the public databases. A large proportion (106 out of 146) of the brayrhizobia strains sequenced were isolated from the nodules of various legumes species. The plants from which the most isolates were sequenced are: (i) *G. max* (36 strains), (ii) *Vigna unguiculata* (12 strains), (iii) *Aeschynomene* spp. (9 strains), (iv) *Phaseolus* spp. (9 strains) and (v) *Lupinus* spp. (7 strains). The other strains originated from various environments (different types of soils, water, biofilm, plant or animal tissue) (Table S1). The majority of bradyrhizobial genomes have sizes ranging from 7 to 10 Mb, with an average of 8.5 Mb. However, a few strains (CCH10-C7, CCH4-A6, NFR13 and SG-6C) have smaller genomes with sizes between 4 and 6 Mb. Among the 33 strains with a complete or near-complete (i.e. <10 contigs) genome sequence, four strains (BTAi1, NK6, DOA9, G22) contain one or several plasmids. Remarkably, strain DOA9 is the only *Bradyrhizobium* strain carrying a symbiotic plasmid harbouring *nod, nif* and T3SS genes [38] and strain NK6 contains four plasmids [39].

To better understand the evolutionary origins of *nod* and T3SS genes in bradyrhizobia, we first aimed at establishing a robust phylogeny of these organisms, which may later be used as a reference to infer evolutionary scenarios for the gain and loss of symbiotic genes. We built a phylogenetic tree based on the concatenated core genome (453 single-copy genes), which we rooted using several representatives from the neighbouring genera *Nitrobacter* and *Rhodopseudomonas* as an outgroup. This analysis yielded six phylogroups within *Bradyrhizobium*, each supported by 100% bootstrap values (Fig. 1). Phylogroup 1 clustered with more than half of the studied genomes. A large proportion of the bradyrhizobia belonging to this phylogroup were isolated from soybean, which is the most important legume crop in the world, thus explaining the over-representation of this phylogroup. Phylogroup 1 gathers strains from the species *Bradyrhizobium japonicum* and *B. diazoefficiens*, as well as strains assigned to various other taxa: *Bradyrhizobium yuanmingense* CCBAU10071^T, *Bradyrhizobium vignae* LMG28791^T, *Bradyrhizobium forestalis* INPA54B^T, *Bradyrhizobium shewense* ERR11^T, *Bradyrhizobium sacchari* BR10280^T, *Bradyrhizobium stylosanthis* BR446^T, *Bradyrhizobium manausense* BR3351^T, *Bradyrhizobium neotropicale* BR10247^T and *Bradyrhizobium centrolobii* BR10245^T. Phylogroup 2 clustered six photosynthetic bradyrhizobia strains, including *Bradyrhizobium oligotrophicum* S58^T, as well as the non-photosynthetic strain STM3843. All seven strains in phylogroup 2 were reported to nodulate some *Aeschynomene* species in an NF-independent manner; they lack *nod* genes, with the exception of strain ORS285, which is able to use both NF-dependent and NF-independent strategies, depending on the host plant [5, 40, 41]. Phylogroup 3 only included two strains, ARR65 and Tv2a-2, which were isolated from nodules of *Stylosanthes viscosa* and *Tachigali versicolor*, respectively [42, 43]. These two strains notably differ from their closest relatives in phylogroup 2 through the presence of *nod* genes. Phylogroup 4 gathered 29 strains from various species, including the type strains *B. elkanii* USDA76^T, *Bradyrhizobium pachyrhizi* PAC48^T, *Bradyrhizobium brasilense* UFLA03-321^T, *Bradyrhizobium viridifuturi* SEMIA690^T, *Bradyrhizobium tropiciagri* SEMIA6148^T, *Bradyrhizobium mercantii* SEMIA6399^T, and *Bradyrhizobium*

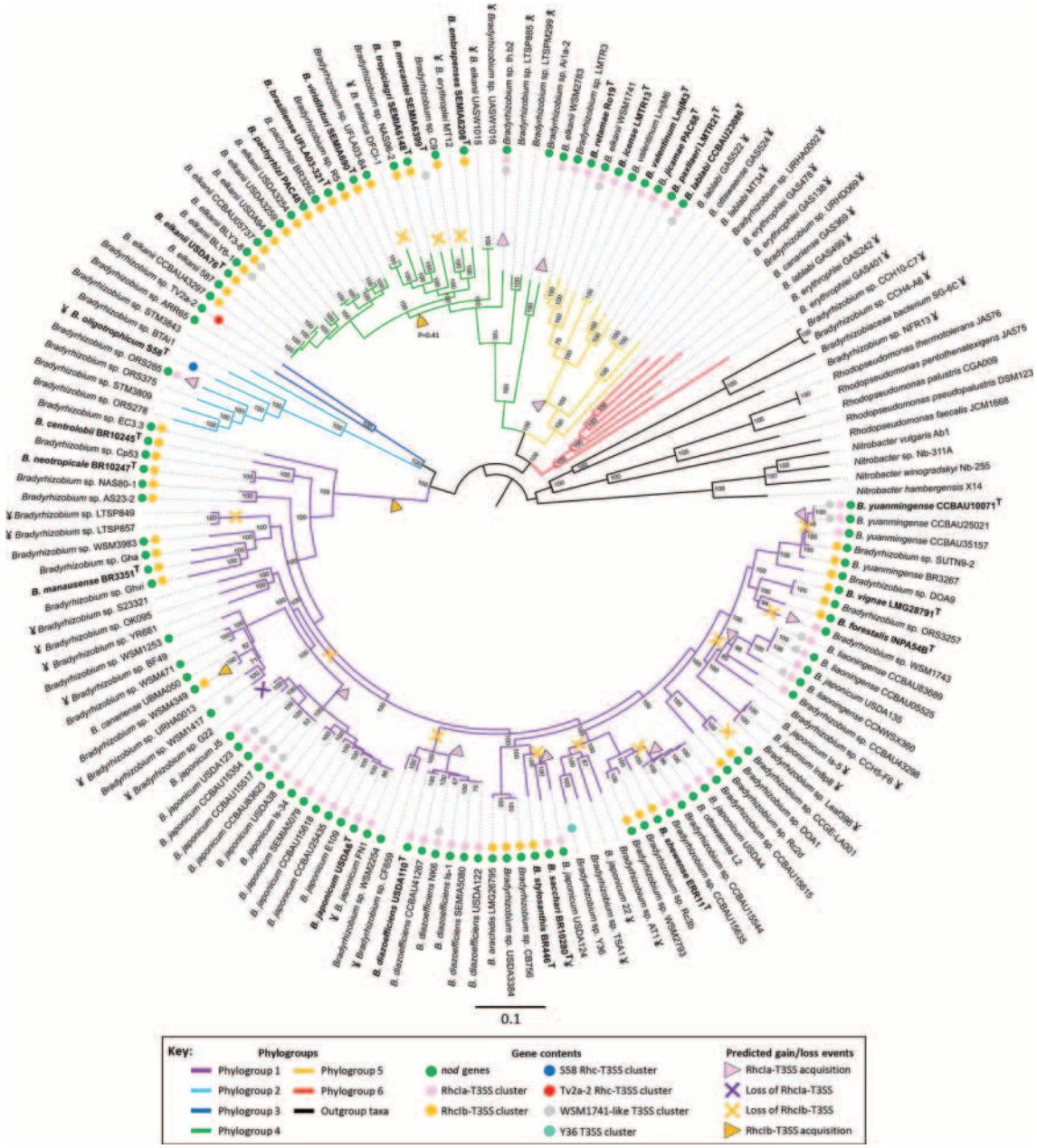


Fig. 1. Phylogenomic tree and RhcI-T3SS cluster gain–loss prediction inferred for the genus *Bradyrhizobium* and its close relatives. The phylogeny was inferred from 453 single-copy core genes for 146 bradyrhizobia strains, and several species of the neighbouring genera *Rhodopseudomonas* and *Nitrobacter*. Node support values are indicated at each branch. The phylogroup to which each strain belongs is indicated by the colour of the branch according to the key provided. The presence of *nod* or T3SS clusters in each strain is shown by colour circles referenced in the key. Probable RhcI-T3SS cluster acquisition and loss events are represented by a colour-coded arrowhead and a cross, respectively, according to the key provided. For all these predicted events, the probability was at least 0.5 except for the gain event of phylogroup 4 for which the probability was 0.41. The scale bar indicates the number of amino acid changes per site. The symbol ¥ denotes the strains that were not isolated from root nodules.

embrapense SEMIA6208^{T}. Similarly, phylogroup 5 gathered 14 strains from several species, such as *Bradyrhizobium retamae* Ro19^{T}, *Bradyrhizobium icense* LMTR13 ^{T}, *Bradyrhizobium valentinum* LmjM3^{T}, *Bradyrhizobium jicamae* PAC68^{T}, *Bradyrhizobium paxllaeri* LMTR21^{T} and *Bradyrhizobium lablabi* CCBAU23086^{T}. Phylogroup 6 contained six strains, all of which were isolated from soils. Finally, the four strains with the smallest genome (CCH10-C7, CCH4-A6, NFR13 and SG-6C) formed a clade that branches with representatives of the genus *Rhodopseudomonas*. These groupings were all supported by whole genome-to-genome distances (Fig. S1).

According to this phylogenetic analysis, the taxonomic classification of several strains needs to be revised. Firstly, the four small-genome strains branched with *Rhodopseudomonas*, which indicates that they do not belong to the genus *Bradyrhizobium* but rather to another genus within the family *Bradyrhizobiaceae*. With this in mind, the genome size of these four strain genomes appears to fit well within the range of sizes observed for *Rhodospeudomonas* (3.7–5.4 Mbp) and *Nitrobacter* (3.6–5.0 Mbp) genomes (NCBI genome database, last accessed September 2019), contradicting the idea that they had undergone some recent genome reduction event. Within the genus *Bradyrhizobium*, several strains do not belong to the species to which they were assigned. This is the case for several strains of *B. japonicum* (USDA4; USDA124; USDA135; 22; is5, in8P8) and *B. elkanii* (UASWS1015; WSM2783; WSM1747), which were found outside of the clade containing the type strain of their respective species. In phylogroup 6, strain GAS499 was wrongly assigned to *B. lablabi*, the type strain of which belongs to another phylogroup. The other strains in this phylogroup are *Bradyrhizobium erythroplei* or *Bradyrhizobium canariense*; the genome sequences of their type strains were not yet available and are not covered by this study, but have representatives in other phylogroups, suggesting that a revision is needed for the taxonomy of either phylogroup 6 strains or the distant strains sharing their name. The incorrect naming of some strains was based on their 16S rRNA and/or the internally transcribed spacer (ITS) sequences, while for the strains belonging to the phylogroup 6 and for WSM2783 and WSM1747 the criteria for their taxonomic classification remain elusive [44–46]. It is now recognized that these two genetic markers are not sufficiently discriminating to distinguish *Bradyrhizobium* strains due to limited sequence polymorphisms in this group. It was therefore proposed to use an alternative molecular method based on multilocus sequence analysis (MSLA) for taxonomic assignments [47]. Our phylogenetic analysis implements a genome-based taxonomy approach, which can be seen as a 'super MSLA', which is not limited to a few partial sequences of housekeeping genes but comprises several hundred conserved genes, as recommended by taxonomic authorities [48, 49].

Notably, during the compilation of this paper, two phylogenomic analyses focusing on the genus *Bradyrhizobium* were published, but they did not include exactly the same taxa as our study. One study was also based on the core genome [50], whereas the other was based on the genome-to-genome BLAST distance phylogeny (GBDP) [51]. The phylogenetic trees of these two studies show a very similar topology to that presented here (Fig. 1), which confirms the robustness of these different phylogenomic analyses. In our tree, the four 'non-*Bradyrhizobium*' small-genome strains form a clade that groups together with the *Rhodopseudomonas* and *Nitrobacter* genomes we used as outgroups. In the study by Avontuur *et al.* [50], which included a wider diversity of outgroup, CCH10-C7 was found to be associated with the genus *Afipia*, while NFR13 belonged to *Tardiphaga*. We explored the phylogenetic position of these strains using the Type Strain Genome Server (https://tygs.dsmz.de/) based on GBDP distances, and this suggested that the strains CCH4-16, CCH10-C7 and SG-6C belong to the genus *Afipia*. A detailed taxonomic study is needed to delineate a formal taxonomic name for these mistyped strains, but this is beyond the scope of this study.

Presence of multiple and diverse T3SS clusters among bradyrhizobia

The *rhcN* gene was used in BLAST searches to survey for the presence of a T3SS among our bradyrhizobial genome collection. This gene encoding an ATPase, which is essential for the function of the secretion machinery, is one of the most conserved core T3SS genes and has often been used for phylogenetic analyses [52, 53].

Among the 146 bradyrhizobial genomes analysed, 96 contain at least 1 *rhcN* homologue and among them, 16 contain 1 additional homologue, and 2 additional homologues are only found in the genome of strain WSM1741 (Fig. 2). This single-gene search analysis suggests that a large proportion (66%) of the bradyrhizobia strains may contains a T3SS gene cluster, and that some of them may even carry several such clusters – a situation already observed in *Ensifer fredii* strains NGR234, HH103 and USDA257, which each contain two T3SS clusters borne on separate plasmids [54–56].

Previous phylogenetic analyses based on various T3SS core proteins, including RhcN, indicated that the T3SS evolved into seven distinct families that spread between Gram-negative bacteria by HGT [52, 57]. So far, the characterized T3SSs from the different rhizobia all clustered in the same family, Rhc, and were further sub-divided into four sub-groups (annotated as α-RhcI to α-Rhc-III and ß-Rhc, which includes the T3SSs from the β-rhizobium *Cupriavidus taiwanensis*) [17]. To better understand how the different bradyrhizobial RhcN homologues are distributed among these different families, an ML tree was constructed including representatives of all the families and sub-groups that were previously defined. The great majority of bradyrhizobial RhcN homologues clustered in sub-group I of the Rhc family (α-RhcI), forming a large clade (annotated A on Fig. 2), which also contains homologues from *Ensifer* and *Mesorhizobium*. Other RhcN homologues seem to be related to the α-RhcI group, but are different from previously defined subgroups [17]; they branch at the base of the α-RhcI group and form a distinct clade (annotated B), grouping genes found in the strains *B. oligotrophicum* S58^T and *Bradyrhizobium* sp. Tv2a-2. Genes from a group of 17 strains, including the type strains *Bradyrhizobium mercantei* SEMIA6399^T, *Bradyrhizobium lablali* CCBAU23086^T and *B. yuanmingense* CCBAU10071^T, form a separate clade (annotated C) that branched next to the Ysc family, which includes proteins from animal pathogenic and symbiotic bacteria. The gene from strain Y36 (annotated D) clearly falls within the Ysc family. When a genome carries several RhcN homologues, at least one copy systematically belongs to the α-RhcI group, with the other copies being associated with clade C, except for strains LmjM3, LmjM6 and WSM1741, whose extra RhcN genes form a clade (annotated E) that is basal to

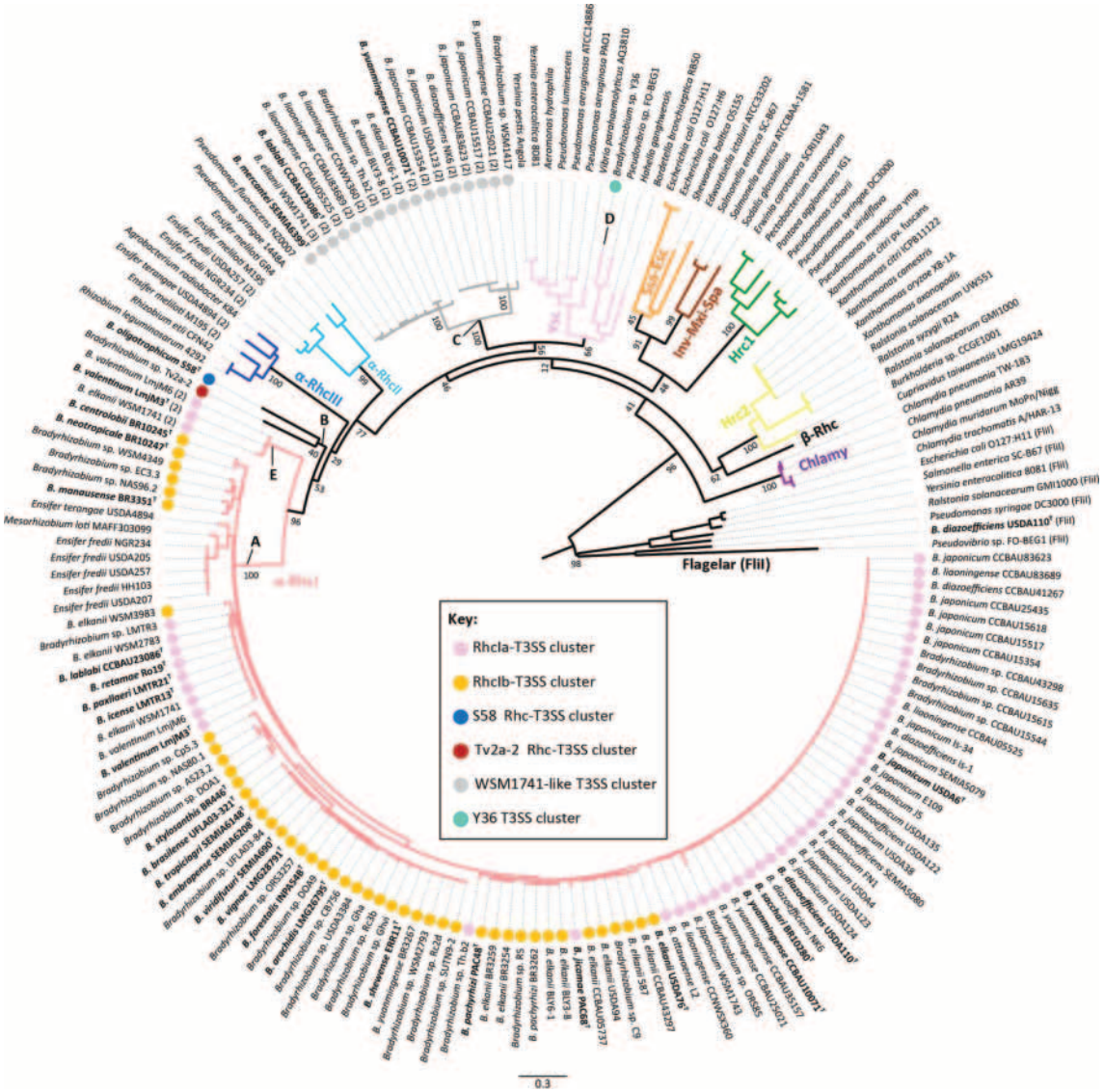


Fig. 2. ML phylogenetic tree of the SctN/RhcN proteins from bradyrhizobia and several representatives of symbiotic, pathogenic and plant-associated bacteria. The ML analysis was performed using mega X and 500 bootstrap replicates. The different SctN families defined according to Troisfontaine and Cornelis [52] are differently coloured. For the *Bradyrhizobium* strains, the different coloured circles (see the key provided) correspond to the distinct genetic organization of the T3SS clusters to which the sctN gene belongs (see Fig. 3). The scale bar indicates the number of amino acid changes per site. The numbers indicated within the parentheses are used to discriminate between the different SctN/RhcN homologues identified in the same strain.

the α-RhcI family. To distinguish the RhcN homologues that do not fall in one of the Rhc families that have already been defined, i.e. those that are related to the Ysc family (clade C and D), we chose to name them SctN, with respect to the unified nomenclature for T3SS proteins (with the Sct prefix referring to secretion) [58].

To further refine the classification of the T3SS systems in bradyrhizobia, we analysed the genetic organization of the gene clusters that encode the T3SS machinery. In total, we identified six different arrangements of T3SS genes among the bradyrhizobial strains (Fig. 3). The majority of T3SS gene clusters belong to the α-RhcI family, with two sub-families named RhcIa and RhcIb, which correspond to the gene clusters already described for *B. diazoefficiens* USDA110^{T} and *B. elkanii* USDA61, respectively (Fig. 3) [9, 59]. In both sub-families, all genes encoding the core components of the T3SS machinery are arranged in three distinct regions (Fig. 3). The main central region (I) contains in the following order the *rhcC1* gene in one direction and the *nopB*, *rhcJ*, *nolU*, *nolV*, *rhcN*, *rhcZ*, *rhcQ*, *rhcR*, *rhcS*, *rhcT* and *rhcU* genes, which are transcribed in the opposite direction. This central region is perfectly conserved in the two sub-families, whereas two additional conserved regions at the extremities of the gene clusters (region II and region III) differ in the details

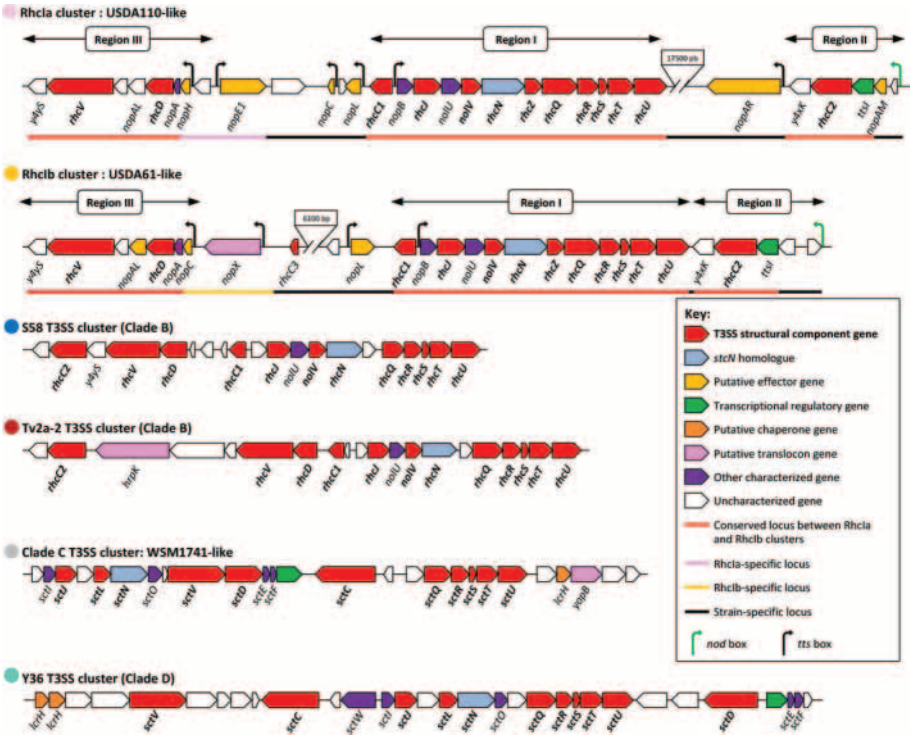


Fig. 3. Genetic organization of the different T3SS gene clusters identified in the genus *Bradyrhizobium*. The six different T3SS gene clusters identified are designated by the same colour circles used in Fig. 2 and the name of a representative strain in which the corresponding T3SS cluster is identified. Open reading frames are represented by coloured arrows according to the key provided and showing the transcription sense. The *nod* and *tts* boxes are represented by green and black arrows, respectively. The two-way arrows above the Rhcla and Rhclb-T3SS clusters represent the three genetic regions defined by Tampakaki [17]. The most common names previously used to characterize the T3SS gene cluster in rhizobia were applied for the T3SS gene cluster found to belong to the Rhc family (according to Fig. 2), while for both the WSM1741-like and Y36-like T3SS clusters, which do not belong to the Rhc family, the *sct* gene nomenclature was used.

of their gene contents. Notably, in region II, the predicted CDSs between the *nod* box and the *ttsI* regulator differ and are strain-specific. In strain USDA110^T, the RhcIa cluster starts with a CDS (bsl1845) encoding a protein of unknown function, followed by the putative effector gene *nopAM* [60], while in USDA61, the RhcIb clusters are predicted to contain two divergently transcribed CDSs encoding hypothetical proteins. Region III starts with two different small genes encoding secreted proteins of unknown function: *nopC* (RhcIb cluster) or *nopH* (RhcIa cluster) [61, 62]. Notably, the region upstream of the RhcIb cluster codes for the NopX protein, which is homologous to the putative translocon HrpF in *Xanthomonas* [63]; the corresponding gene is absent in all the strains with a RhcIa cluster, such as strain USDA110^T.

One of these two gene arrangements is systematically found in all the strains containing a *rhcN* gene of the α-RhcI family, but some minor variations occur among strains: (i) gene predictions in the region upstream of the *ttsI* gene are often ambiguous, sometimes with a prediction of a CDS showing some similarities with the *nopAM* gene identified in strain USDA110^T; (ii) a rearrangement of the region II provoked by transposase insertions and leading to a separation of the *ttsI* and *rhcC2* genes (observed only in ORS285 strain); and (iii) the occasional presence of an intergenic region of around 650 bp between *rhcS* and *rhcT*. Interestingly, this 'intergenic' region is predicted to contain a *tts* box, suggesting that the genes of such a region I might be transcribed in two distinct operons.

Notably, the T3SS gene clusters from clade E share the same genetic organization as RhcIa and should therefore be considered to be a bona fide RhcI-T3SS cluster. In contrast, T3SS gene clusters from strains S58^T and Tv2a-2 (clade B) both have a different genetic organization than RhcI types and should therefore constitute new types of Rhc cluster (Figs 2 and 3). In particular, the T3SS cluster in the Tv2a-2 strain harbours a homologue of the *hrpK* gene, which is found in some non-canonical T3SS gene clusters from *Ensifer* strains and phytopathogenic bacteria, where it is thought to encode a translocon component [13, 17].

Analysis of the regions surrounding the *rhcN/sctN* homologues of strain Y36 (clade D) or of the strains belonging to clade C revealed additional types of T3SS gene clusters, suggesting a distinct origin of these unconventional T3SSs (Figs 2 and 3). Determining whether these atypical T3SS clusters are functional and can affect the symbiotic properties of the strains requires further investigation.

A striking difference of the T3SS clusters from clades B to D in comparison to the RhcI clusters is the absence of the *ttsI* regulatory gene. The fact that *ttsI* is under the control of the master regulator NodD [18, 19, 64] supports the hypothesis that the ancestor of the RhcI-T3SS cluster was formed in the presence of the *nod* gene cluster, and that later propagation of the T3SS would rely on its co-transfer with *nod* genes, most probably via a symbiotic island gathering these two gene clusters.

The *nod* and T3SS genes share a common history in the genus *Bradyrhizobium*

To test this hypothesis, we scrutinized the genomes of the 146 *Bradyrhizobium* strains for the presence of *nodABC* genes using BLASTN and observed that 100 strains contained these canonical *nod* genes (Fig. 1). Strikingly, more than 90% of them (92 out of the 100) also contain the RhcI-T3SS cluster. In fact, all the strains containing a RhcI-T3SS cluster contain the *nodABC* genes, except for strain *Bradyrhizobium liaoningense* CCBAU83689 (Fig. 1). We then examined the genome of the 33 strains with a complete or near-complete sequence to see whether the *nod* and RhcI-T3SS gene clusters are co-localized or not. For the 12 strains containing a RhcI-T3SS cluster, this cluster is systematically found inside the symbiotic island containing the *nod* and *nif* genes (Table S3). Conversely, for the three strains (S58^T, NK6 and WSM1417) with an atypical T3SS cluster, this cluster is not found on the symbiotic island but on another horizontal acquired island for S58^T and WSM1417 strains, or on a plasmid for NK6 strain (Table S3).

To verify that the RhcI-T3SS and *nod* gene clusters follow the same evolutionary history, we compared the phylogenetic trees based on the concatenation of the NodA, NodB and NodC proteins or on the concatenation of the RhcC2, RhcN and RhcV proteins — each corresponding gene belonging to one of the three regions composing the T3SS cluster. Several incongruences occur between the trees, which likely result from ancient recombination events between symbiotic islands, as may have occurred after symbiotic islands from different origins were acquired independently and co-inhabited in the same rhizobial genome. This hypothesis of recombination among clusters is further supported by the evolutionary scenarios that we estimated for component genes of these clusters. Inferred scenarios show that most HGT events were inferred uniquely in each of the *nod* or *rhc* gene trees, indicating that independent importation of foreign sequences at each gene locus occurred regularly over the evolution of the bradyrhizobial symbiotic islands (online Supporting Data, http://dx.doi.org/10.6084/m9.figshare.12191103). The (rare) observation of two RhcI-T3SS clusters with divergent *rhcN* genes within the extant genomes of strains WSM1741, LmjM3 and LmjM6 finally proves the possibility of a co-habitation being required for such a recombination event (Fig. 2).

Nevertheless, the phylogenetic trees for NodABC and RhcC2VN present many matching clades grouping corresponding sequences from the same strains (Fig. 4), supporting the hypothesis of co-evolution between these two gene clusters, at least in recent times.

We thus aimed to quantify the co-evolution of single-gene families within and between the *nod* or RhcI-T3SS gene clusters. To do so, we estimated evolution scenarios describing the events of vertical and horizontal dissemination that marked the diversification of key genes present in these gene clusters. We then computed pairwise co-evolution scores based on how much each pair of gene families had shared events in their scenarios. As expected, genes show the strongest co-evolution association with those that are co-located with them in modern genomes: genes within the *nod* operon (i.e. *nodA*, *nodB* and *nodC*) and those within the Rhc-T3SS gene cluster (i.e. *rhcC2*, *rhcD*, *rhcN*, *rhcV* and *nopC*) each form modules of genes with significantly strong co-evolution history. Co-evolution between the *nod* and RhcI-T3SS modules was also strong and significant, even though slightly lower than within these modules (Fig. S3, http://dx.doi.org/10.6084/m9.figshare.12191103). This is evident when considering all shared evolutionary events, including both vertical and horizontal modes of co-evolution (Fig. S3 a, c), or when considering only co-transfer events, i.e. only events of joint acquisitions via HGT (Fig. S3b, d); this suggest that these genes were often co-acquired and subsequently conserved together in genomes.

Taken together, these data support the hypothesis that bradyrhizobia gained *nod* and RhcI-T3SS gene clusters simultaneously through the acquisition of a symbiotic island, while the atypical T3SS clusters were gained independently. The subsequent vertical inheritance and preservation of the RhcI-T3SS in most of the strains containing *nod* genes – notwithstanding a few recombination events – also indicates that this secretory machinery plays an important role in their symbiotic lifestyle.

T3SS was gained and lost multiple times in the genus *Bradyrhizobium*

The distribution of the RhcIa and RhcIb clusters is not random within the genus *Bradyrhizobium*, with the cluster types occurring in homogeneous patches within clades of the core genome tree (Fig. 1). The distribution of the RhcI-T3SS cluster types is not uniform with regard to their host genome phylogroup: (i) strains belonging to phylogroups 6 and 2 (with the exception of the strain ORS285), as well as the two members of phylogroup 3 have no RhcI cluster; (ii) phylogroup 5 strains possessing a T3SS all have the RhcIa cluster type; (iii) all phylogroup 4 strains possessing a T3SS, except strain th.b2, have the RhcIb cluster; (iv) RhcIa and RhcIb clusters are irregularly distributed within phylogroup 1.

To determine the origins of this heterogeneous distribution of RhcIa and RhcIb cluster types, we first explored the evolution scenarios computed for each component gene. However, each gene's sample of scenarios presented a set of many probable events, distributed over the reference *Bradyrhizobium* tree, and while statistical agreement between those sets can be computed (see the co-evolution scores above), it is difficult

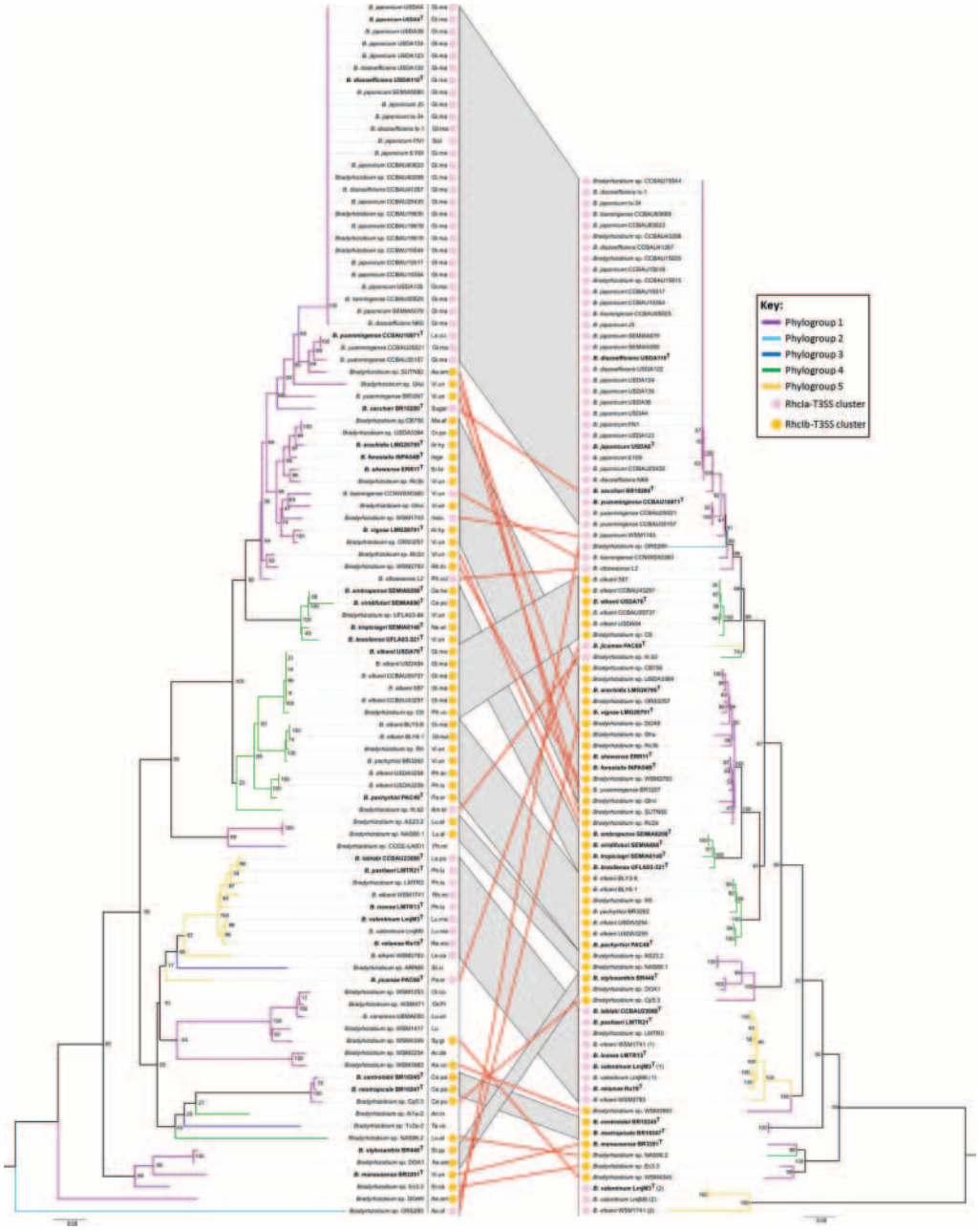


Fig. 4. Comparison of phylogenetic trees for concatenated NodABC (left) and RhcC2NV (right) proteins from *Bradyrhizobium*. The two ML phylogenetic trees were constructed using MEGA X and 500 bootstrap replicates. *Bradyrhizobium* sp. UFLA03-84 is not included in the RhcI-T3SS tree, since this strain lacks the region III containing the *rhcV* gene. The scale bar represents the number of amino acid substitutions per site. The colouring scheme for tree branches and circles is indicated in the key provided. The phylogenetic congruence between the two trees is represented by grey bands linking the different clades, while incongruence is represented by red lines. The environment for isolation of the strains is indicated in the field next to the strain names on the left-hand tree *Acacia dealbata* (Ac.de), *Aeschynomene afraspera* (Ae.af), *Aeschynomene americana* (Ae.am), *Amphicarpaea bracteata* (Am.br), *Andira inermis* (An.in), *Arachis hypogaea* (Ar.hy), *Centrolobium paraense* (Ce.pa), *Centrosema pubescens* (Ce.pu), *Crotalaria paulina* (Cr.pa), *Desmodium heterocarpon* (De.he), *Erythrina brucei* (Er.br), *Erythrina costaricensis* (Er.co), *G. max* (Gl.ma), *Indigofera* sp. (Indo.), *Inga* sp. (In.sp), *Kennedia coccinea* (Ke.co), *Lablab purpureus* (La.pu), *Leobordea carinata* (Le.ca), *Lespedeza cuneata* (Le.cu), *Lupinus albescens* (Lu.al), *Lupinus angustifolius* (Lu.an), *Lupinus maria-josephae* (Lu.ma), *Lupinus* sp. (Lu.sp), *Macrotyloma africanicus* (Ma.af), *Neonotonia wighti* (Ne.wi), *Ornithopus compressus* (Or.co), *Ornithopus pinnatus* (Or.pi), *Pachyrhizus erosus* (Pa.er), *Phaseolus acutifolius* (Ph.ac), *Phaseolus lunatus* (Ph.lu), *Phaseolus microcarpus* (Ph.mi), *Phaseolus vulgaris* (Ph.vu), *Retama monosperma* (Re.mo), *Rhynchosia minima* (Rh.mi), *Rhynchosia totta* (Rh.to), Sugar cane (Su.ca), *Stylosanthes guianensis* (St.gu), Stylosanthes viscosa (St.vi), *Syrmatium glabrum* (Sy.gl), *Tachigali versicolor* (Ta.ve) and *Vigna unguiculata* (Vi.un).

to summarize that information in a consensus scenario for the whole gene cluster (see high-probability co-events in the online Supporting Data; http://dx.doi.org/10.6084/m9.figshare.12191103). For greater ease of interpretation, we thus chose to complement our analysis with a simpler method of inference of the history of events of gain and loss of Rhc clusters. We used the program Gloome, which relies on the profile of presence and absence of the whole RhcI clusters in genomes and performs a stochastic mapping of events over the reference tree, to estimate a single scenario of evolution for each of the RhcI cluster types along the *Bradyrhizobium* tree [34]. According to the scenario estimated with Gloome, an equal number of gain and loss events (14 each) occurred across the genus tree, but with different frequencies depending on the phylogroup.

A few seminal events of independent acquisition of the RhcI-T3SS cluster define the main structure of their distribution: the acquisition of the RhcIa cluster in the ancestor of the phylogroup 1 clade corresponding to the species *B. japonicum* and *B. diazoefficiens*; another acquisition of the RhcIa cluster in the ancestor of the phylogroup 5 clade regrouping *B. lablabi*, *B. paxllaeri*, *B. jicamae*, *B. valentinum*, *B. icense* and *B. retamae*; and the acquisition of the RhcIb cluster in the ancestor of the phylogroup 4 clade regrouping the species *B. elkanii, Bradyrhizobium pachirizi*, *Bradyrhizobium braziliense*, *B. viridifuturi*, *Bradyrhizobium tropicagri*, *B. mercanti* and *Bradyrhizobium embrapenses*. These seminal horizontal transfer events were followed by the conservation of the acquired cluster, leaving patterns of locally homogeneous cluster distribution, suggesting that strong purifying selection acted on these acquired genes. Another acquisition of the RhcIb cluster was inferred at the root of the phylogroup 1, but this event was later followed by independent losses of the cluster: it was lost by the ancestor of the group formed by *B. japonicum* and unclassified *Bradyrhizobium* strains such as WSM1417 and WSM1253. The ancestor of *B. japonicum* later regained a T3SS cluster of the type RhcIa. Other loss events are associated with gains of the other cluster type on the same branch of the tree, and six such replacements of the RhcIb cluster by the RhcIa type occurred, all within phylogroup 1 (Fig. 1).

This scenario of repeated replacements is reflected in the phylogeny of the RhcC2VN proteins, where clades of Rhc proteins associated with the RhcIa organization recurrently emerge from a larger clade of proteins associated mostly with the RhcIb type (Fig. 4). The RhcIa clade at the top of Fig. 4 regroups many almost identical sequences, suggesting a more recent spread of this cluster type, with repeated events of horizontal dissemination into distinct clades of phylogroup 1. Interestingly, most of these RhcIa phylogroup 1 strains were isolated from *G. max* nodules. It is very likely that a strong selective pressure, probably driven by the interaction with the plant, has favoured the exchange of RhcI-T3SS clusters within this phylogroup. The selective advantage of switching to a RhcIa cluster type may be specific to phylogroup 1, as *B. elkanii* strains of phylogroup 4 also isolated from *G. max* nodules all harbour a RhcIb cluster.

The tree based on concatenated RhcC2VN proteins presents the evolutionary point of view of the genes encoded in the cluster (Fig. 4), which shows that the RhcIa type evolved twice from a clade associated with the RhcIb type. These events correspond to a gain by the ancestor of *B. japonicum* (phylogroup 1), and another gain in the ancestor of phylogroup 5 Rhc+ strains (Fig. 1). The lineages of Rhc proteins involved in these two gain events are not related, with those found in phylogroup 5 branching deeply in the RhcC2VN phylogenetic tree. Interestingly, the diversity of proteins within this deep clade mirrors that of their host genomes. This suggests that in phylogroup 5, the *rhc* genes (or at least large segments of the locus including the *rhcC2*, *rhcV* and *rhcN* genes) were transmitted vertically after the gain of the ancestral RhcIa type cluster. This unique gain event in phylogroup 5 is confirmed by reconciliation-based scenarios (see ML reconciliations in the Online Supporting Data; http://dx.doi.org/10.6084/m9.figshare.12191103), but those scenarios additionally implied that it was followed by occasional events of homologous recombination of individual *rhc* genes among phylogroup 5 strains.

To evaluate whether these repeated changes of RhcI type were regularly associated with such a reshuffling of the alleles of core *rhc* genes, we then studied the individual trees of genes located in different regions of the RhcI locus: *rhcV* in region III, *rhcN* in region I and *rhcC2* in region II (Figs 3 and S2). The evolutionary histories of these genes appear to be mostly congruent, as they share the majority of ancient splits, including those leading to the emergence of the clades that are homogeneous for their RhcI type. This strongly indicates that the whole RhcI-T3SS cluster locus mostly evolved as a linked genetic unit, being transferred from one genomic background to another as a single fragment. In addition, it shows that the diversification of *rhc* genes is not associated with switches in RhcI cluster type; it is only the genetic organization of the locus that changed when changes of RhcI type occurred. Each switch of cluster type could in fact have been mediated through a single homologous recombination anchored in the conserved regions, from upstream of *nopA* to downstream of *rhcC1* (Fig. 3, red underlined regions), converting the type-specific region located in between (Fig. 3, gold and pink underlined regions). This single-step scenario for conversion of a RhcIb-type cluster into a RhcIa cluster, and vice versa, makes it a potentially frequent event.

Indeed, a few RhcI gain and loss events were observed in recent lineages, mostly limited to single genomes in our sample. These recent changes may represent random flickers in the evolutionary history with no particular selective value, or could alternatively be the sign of positive selection for the acquisition of the T3SS, or relaxation of purifying selection leading to its loss. The type-specific variant regions are also the site of the accumulation of strain-specific genetic material – mostly insertion sequences (Fig. 3, black underlined regions) – indicating that these regions are foci of rapid gene content evolution, and that strain-specific genetic apparatus have not yet undergone strong selective pressure.

On a longer evolutionary timescale, the fact that sporadic T3SS acquisitions (of either RhcIa or RhcIb type) always occurred in a Nod+ background suggests that the T3SS may provide some symbiotic function that is under positive selection in strains that already produce the Nod signal. Similarly, sporadic losses occurred in some lineages or recent clades, including independent losses in strains DFCI-1, SEMIA6399^T and MT12 in phylogroup 4, and a loss in the ancestor of strains LTSP849 and LTSP857 in phylogroup 1. Notably, in all these cases, the loss of the T3SS is correlated with the loss of the *nod* genes on the same branch. This suggests a scenario where loss of the *nod* genes led to a relaxation of the purifying selection acting on the existing RhcI-T3SS cluster, which finally resulted in its loss. However, another explanation (that does not involve natural selection) is that the *nod* and *rhc* genes could have been acquired concomitantly or lost via homologous recombination between the surrounding core genes, thus resulting in insertion or deletion of the whole symbiotic island.

Altogether, these data indicate that the T3SS have been acquired a few times by ancestors of bradyrhizobial lineages and that the distribution of the T3SS types was later disturbed by a dynamic history of gain via horizontal gene transfer and loss, perhaps driven by selective pressure related to symbiosis with the plant host.

Finally, we also analysed the phylogenetic distribution of the atypical T3SS clusters and observed that these clusters are distributed patchily among phylogroups 1 to 5 and absent in phylogroup 6 (see Fig. 1). The patterns of conservation of atypical clusters among some groups of phylogenetically related strains within the *B. japonicum*, *B. lianoningense*, *B. yuanmingense* and *B. elkanii* species suggest separate events of acquisition in their respective common ancestors. Other than that, no correlation can be drawn between the presence of one of these atypical T3SS cluster and any other symbiotic or lifestyle features.

Possible new cases of *Bradyrhizobium* strains nodulating through their T3SS in the absence of NF

Considering the environment of isolation of the strains, it appears that 38 out of the 40 strains isolated from soils or other environments, but not nodules, lack both the *nod* and RhcI-T3SS genes. The only two exceptions are *B. sacchari* BR10280^T and *B. japonicum* FN1, isolated respectively from sugarcane roots and from Manitoba soils. These 38 strains are distributed in 5 out the 6 phylogroups and notably include all the 6 representative strains of phylogroup 6 represented in our dataset (Fig. 1). This suggests that the presence of the symbiotic genes could have a negative impact on bacterial fitness in the absence of an interaction with a plant, and be counter-selected outside of nodules. Furthermore, the strains from the divergent phylogroup six may not have a genomic background allowing them to nodulate legume plants upon acquisition of a symbiotic island.

Conversely, most isolates from plant nodules do contain *nodABC* genes, which reinforces the idea that the NF-dependent symbiotic process is the predominant type of interaction, even though some exceptions (8 isolates out of 106) exist. As previously indicated, five strains that were isolated from *Aeschynomene* species (*A. indica* and *A. evenia*) and lack *nod* and RhcI-T3SS clusters nodulate *Aeschynomene* species in an NF-independent process [5, 40, 41]. More surprisingly, three strains (*B. mercantei* SEMIA6399^T, *B. liaoningense* CCBAU83689 and *Bradyrhizobium* sp. Y36) isolated from plant species that do not belong to the genus *Aeschynomene* were also found to lack *nod* genes. The strain *B. mercantei* SEMIA6399^T was isolated from an effective nodule of the tropical legume *Deguelia costata* (syn. *Lonchocarpus costatus*), an important plant species belonging to the tribe *Millettieae* that is native to eastern Brazil [65, 66]. Interestingly, this strain, which is phylogenetically distant to the phylogroup 2 bradyrhizobia nodulating *Aeschynomene,* was confirmed to form effective nodules on its plant of origin but also on *Macroptilium atropurpureum* [65]. This strain is even used as a commercial inoculant for *D. costata*. If the symbiotic properties of the sequenced strain are confirmed, this strain would be the first example of a *Bradyrhizobium* outside phylogroup 2 that is able to use an NF-independent pathway to naturally nodulate legumes other than *Aeschynomene* spp. Strain SEMIA6399^T also lacks a RhcI-T3SS cluster but contains an atypical T3SS cluster. It would be interesting to elucidate whether this T3SS plays a role in the nodulation abilities of this bacterium.

The two others nod-lacking strains (*B. liaoningense* CCBAU83689 and *Bradyrhizobium* sp. Y36) were isolated from *G. max* and from *Phaseolus vulgaris*, respectively, two legume species known to be nodulated by rhizobia with *nod* genes. One may wonder whether these two strains were opportunist bacteria that took advantage of rhizobia excreting NFs to co-infect the same nodules. It would be interesting to verify the ability of these strains to nodulate their original host plants in an NF-independent manner. As previously stated, the strain *B. liaoningense* CCBAU83689 is the only bradyrhizobium that has a RhcI-T3SS cluster but lacks *nod* genes. We can therefore assume that this strain has previously acquired a symbiotic island containing both *nod* and T3SS genes and subsequently lost the *nod* genes but maintained the RhcI-T3SS cluster. While this strain indeed nodulates *G. max,* it remains to be determined whether its RhcI-T3SS is the major determinant governing the nodulation. Similarly, strain Y36 also contains an atypical T3SS cluster whose symbiotic role cannot be excluded.

The absence of *nopX* is correlated with the presence of *nopH* and *nopE*

The translocator proteins form a pore (called a translocon) in the host membrane that allows the passage of effectors into the eukaryotic host cell [67]. The components forming the translocon in rhizobial T3SS remain unknown, but the homology of NopX with the putative translocator HrpF from *Xanthomonas* species suggests that it could be a translocon component [68]. The absence of a gene encoding a translocator protein in all the strains containing a RhcIa cluster

is enigmatic and raises the question of which gene(s) may play this function. A similar case has been reported in the genus *Xanthomonas*, where members of the ancestral, early-branching clade lack *hrpF* [69]. Comparative genomics permitted us to identify specifically in this group of strains two candidate translocator genes (*hpaT* and *hpaH*) for which a role in effector translocation has been subsequently confirmed [69]. Following this approach, we speculated that all the RhcIa+ strains should contain (a) specific gene(s) that could functionally substitute *nopX*. Comparison of the bradyrhizobial genomes highlighted two genes that were specifically associated with the presence of a RhcIa cluster (Figs 3 and 5): *nopH*, which encodes a protein of unknown function of about 100 amino acids, and *nopE*, which encodes a protein of about 500 amino acids that is assumed to be a translocated effector [70]. These two genes are found in the same genetic organization and closely associated with the core T3SS gene clusters in all the RhcIa+ strains; *nopH* is always found upstream of *nopA,* and *nopE* is always found upstream of region III in a genetic organization that is very similar to *nopX* in the RhcIb cluster (Fig. 3).

To the best of our knowledge, the function of NopH is not known, except that this protein was found in the secretome of USDA110^T [61]. In contrast, NopE has been functionally studied, in particular in strain USDA110^T, which harbours two NopE homologues. It has been shown using CyaA reporter fusion that both NopE variants were translocated into plant cells and that they are most probably functionally redundant [70]. Indeed, only a *nopE* double mutant is significantly (positively) affected in its ability to nodulate *Vigna radiata*, suggesting that NopE proteins or a process relying on NopE activity play a negative role in the nodulation of this plant species [70].

NopH and NopE display no sequence similarity with NopX or with known translocators from other bacteria. However, the translocon proteins are not conserved among different bacterial species [71]. The specific presence of *nopH* and *nopE* only in the RhcIa+ strains make them attractive candidates for functional studies of the T3SS. It would be interesting to determine if they play a role in the translocation of effectors, keeping in mind that the functional features already reported for NopE are not incompatible with such a role.

Drastic variation between the effectome of *Bradyrhizobium* strains

In order to determine the content and diversity of T3Es in the genus *Bradyrhizobium*, we screened for the presence of a set of T3Es in the predicted proteomes of all strains using BLASTP (% of identity ≥40% over ≥70% of the length of the protein). We restricted this analysis to rhizobial effectors that have been demonstrated to be secreted through the T3SS and/or shown to affect the symbiosis positively or negatively (Table S2). As shown in Fig. 5 and as expected from previous results, T3E candidates were generally not found in genomes that do not harbour a RhcI-T3SS cluster, except for six strains (CCGE-LA001, WSM2254, ARR65, Tv2a-2, SEMIA6399^T and Ai1a-2). Two of these strains (SEMIA6399^T and Tv2a-2) present an atypical T3SS cluster, so it cannot be excluded that the identified effectors are translocated by the secretory machinery encoded by these atypical T3SS gene clusters. In addition, these two strains and strain Ai1a-2 present a few genes encoding components of the RhcI-T3SS in their genomes, which raises the possibility that their ancestor's genome once harboured a RhcI-T3SS cluster with its cortege of T3Es but lost most of the T3SS genes since (Fig. 2).

For all the strains containing a RhcI-T3SS cluster, the number of identified T3Es is highly variable, ranging from 2 for WSM3983 to 24 for *B. diazoefficiens* USDA122 (Fig. 5). Notably, *B. diazoefficiens* and *B. japonicum* strains isolated from soybean nodules have the largest set of effectors, with an average of 20 T3Es. Nevertheless, this number of T3Es also varies considerably within these 2 species, from 14 to 24. The repertoire of effectors, also called the effectome, is frequently remodelled by the acquisition and loss of genes, a process most likely shaped by the selective pressure of the host plant.

It is worthy of note that there is not a single T3E that is found in all strains with a RhcI-T3SS cluster. While no core effectome can therefore be defined within the genus *Bradyrhizobium*, four T3Es (NopM, NopP, NopT, NopC) are nevertheless present in a majority of genomes and are often found in multiple copies. Interestingly, these four T3Es are also conserved in other rhizobial genera [60], which leads to the speculation that they were part of an 'early' effectome that was acquired horizontally by the ancestors of bradyrhizobial lineages, possibly at the same time as the RhcI-T3SS clusters, and later lost in some specific lineages, perhaps following different selective pressures.

This hypothesis is in line with our reconciliation-based inference of evolutionary scenarios for the T3E genes *nopC*, *nopM*, *nopL* and *nopT*, and with our estimation of how much they co-evolved between themselves and with genes from the *nod* and RhcI-T3SS clusters (Fig. S3). As expected, the gene *nopC* was found to co-evolve strongly with the *rhc* genes encoded within the same locus (for RhcI-type T3SS clusters). Surprisingly, *nopM*, *nopL* and *nopT* formed a co-evolving gene module, even though they are encoded at distant loci. These three T3Es also showed some association with *nod* and *rhc* genes. Specifically, these T3Es are only loosely associated – sometimes not significantly – with either *nod* or *rhc* gene modules when considering both vertical and horizontal modes of co-evolution (Fig. S3a, c), whereas they showed a strong association with them (and *nod* genes in particular) when considering only horizontal co-evolution (Fig. S3b, d). A possible interpretation of these patterns is that these three T3Es were frequently co-acquired with *nod* genes (and with *rhc* genes to a lesser extent), but would often have been lost afterwards, leading to rare vertical co-transmission with either *nod* or *rhc* genes. This supports the 'early effectome' hypothesis, where these T3Es were acquired by bradyrhizobial ancestors as part of the same 'informational package' containing the *nod* and *rhc* genes (possibly on the same symbiotic island) but were

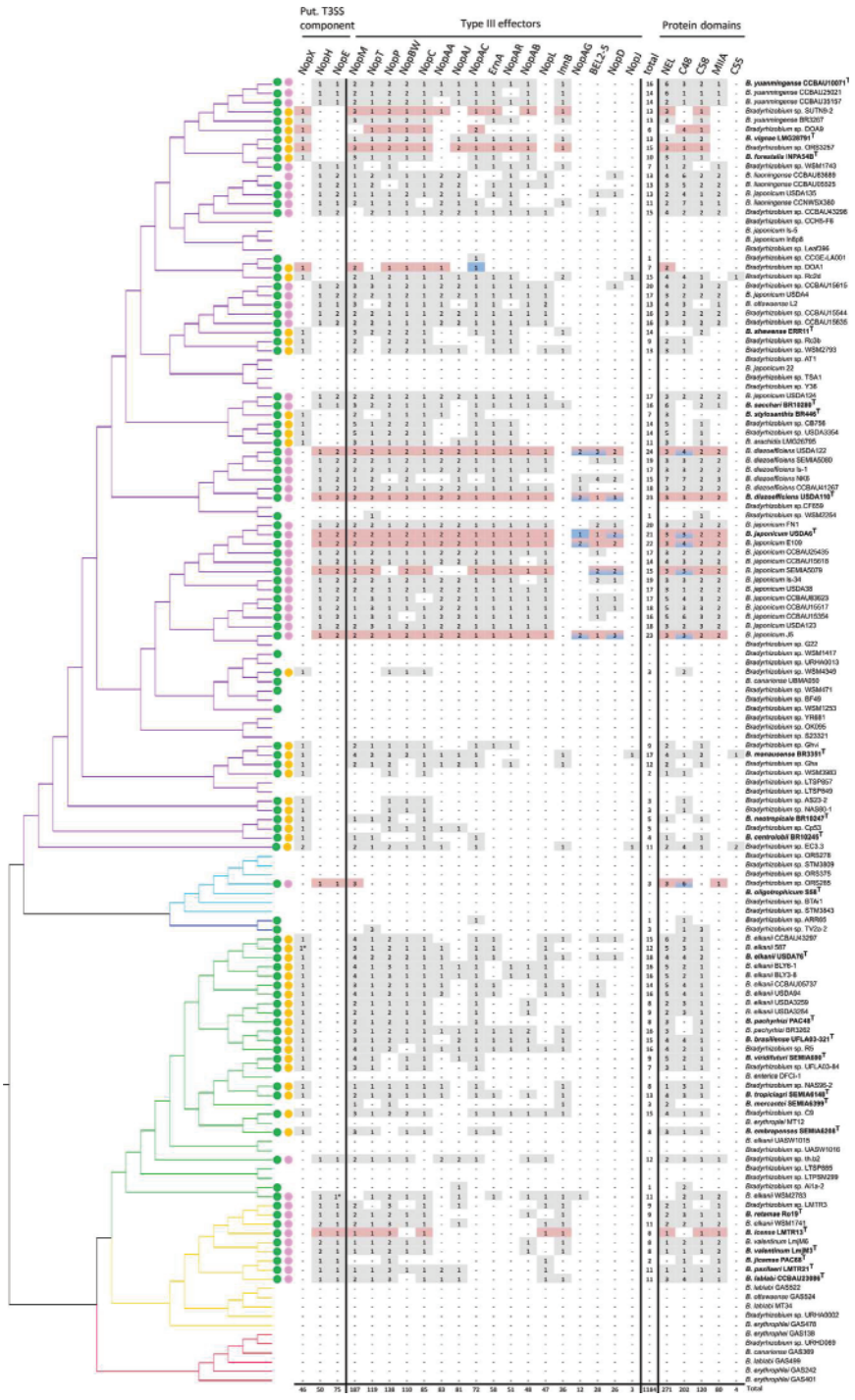


Fig. 5. Representation of the effector candidate repertoire in the genus *Bradyrhizobium*. The number and distribution of some T3SS components, some known Nops and some proteins containing known functional domains identified in some rhizobia effectors are shown. Numbers represent the number of proteins found in each strain according the research parameters defined in the Materials section. NEL, novel E3-ubiquitine ligase. MIIA, metal ion-inducible autocleavage protein. C55, YopJ-type protein. C48, SUMO protease. C58, YopT-type protein. The asterisk (*) indicates frameshift mutations in the corresponding gene. The colouring scheme of Fig. 1 is used to indicate which phylogroup the strain belongs to and the coloured circles indicate the presence of a *nod* gene cluster (in green), a RhcIa-T3SS cluster (in pink) or a RhcIb-T3SS cluster (in yellow). For the 12 strains containing a RhcI-T3SS cluster and for which the sequence genome is complete or almost complete, the boxes containing the number are respectively highlighted in red or blue when the genes are on the symbiotic island or another horizontal acquired region. When the cases are bicoloured, this indicates that different effector homologues can be found on both the symbiotic island and on another horizontal acquired region.

differentially lost or retained in descendant lineages, owing to varied selective pressures.

We also examined whether the distribution of any T3E was associated with the presence of the RhcIa or the RhcIb cluster types. No strict correlation except those discussed before for NopX, NopH and NopE was observed, indicating that the ancestral RhcIa or RhcIb clusters most probably shared similar T3E repertoires. However, at a lower frequency, two other effectors, NopAG and NopJ, both encoding a putative acetyltransferase, were found to be specifically associated with the RhcIa and RhcIb cluster, respectively. NopAG is present in seven strains of *B. japonicum* and *B. diazoefficiens* and the distant strain WSM2783, with no homologue being found outside the genus *Bradyrhizobium*. This exclusive distribution suggests that this T3E has been specifically acquired or evolved within the *B. diazoefficiens*/*B. japonicum* group and later horizontally transferred to the strain WSM2793. The NopJ effector was only found in three strains that are phylogenetically very distant from each other. This T3E, which belongs to the YopJ effector family in pathogenic bacteria, had previously only been reported in one rhizobial genome, *Ensifer fredii* strain NGR234 [60, 72]. It is therefore possible that NGR234 and these three *Bradyrhizobium* strains acquired NopJ independently by lateral gene transfer.

For the 12 strains containing a RhcI-T3SS cluster and for which the sequence genome is complete or almost complete, we have also examined the localization on the genome of the predicted T3Es. As shown in Fig. 5, most of the T3Es are associated with the symbiotic island. The only exceptions are observed for homologues of *nopAG*, *nopD*, *bel2-5* or other genes encoding C48 SUMO proteases, which are found outside of the symbiotic island. However, a *tts* box can be identified upstream of most of these genes, reinforcing the idea that they would be bona fide T3Es. Furthermore, these genes are also found within regions with low GC content and different codon usage with respect to the rest of the chromosome. This suggests that their localization away from symbiotic genes could be the result of a fragmentation of the initially acquired symbiotic island (such as has been described for the NK6 strain [39]) or from another horizontal acquisition event.

The diversity and the effectors number are most probably underestimated

T3Es are often modular proteins constituted of multiple conserved domains [73]. This particular architecture complicates the search for homologues because of the low level of similarity between effector proteins outside of these conserved domains. Thus, the search for specific functional domains has been proposed as a more sensitive approach to identify new candidate effector proteins [74]. Among the rhizobial effectors, five functional domains have been identified: two eukaryotic domains, NEL (for novel E3 ligase/IPR029487) identified in NopM [75, 76] and the SUMO protease domain (peptidase family C48/IPR003653) identified in NopD and BEL2-5 [77, 78]; the YopT-type domain (peptidase family C58/IPR006473) identified in NopT [79, 80], the YopJ-type domain (peptidase family C55/IPR005083) identified in NopJ [72] and the recent identified MIIA domain (metal ion-inducible autocleavage/IPR011086) identified in NopE [81]. In order to highlight new putative T3Es, we performed an InterProScan analysis on the annotated bradyrhizobial genomes using the MaGe interface on the MicroScope platform (Fig. 5). All proteins identified as containing one of these domains were found specifically in strains that have a RhcI-T3SS cluster or in five of the six strains (WSM2254, ARR65, Tv2a-2, SEMIA6399^T and Ai1a-2) that were discussed above and are supposed to have lost their T3SS. The search for YopT and YopJ-type domains as well as for the MIIA domain retrieved 120, 4 and 80 proteins, respectively, which correlates well with the number of homologues identified by BLASTP for NopT (119 homologues), NopJ (three homologues) and NopE (75 homologues).

In contrast, the search for the NEL and the SUMO protease domains returned a far higher number of putative T3Es (271 and 202 proteins) than the BLASTP search for homologues of NopM protein or the two SUMO proteases NopD and BEL2-5, which retrieved 187, 28 and 26 homologous proteins, respectively. We thus identified 84 additional NopM-family T3Es and 180 additional proteins containing a SUMO protease domain. This suggests that there are still many T3Es to be discovered in the genus *Bradyrhizobium*, some of which may play an important symbiotic role.

Conclusion

The importance of the T3SS in the symbiotic ability of rhizobia remains unclear. In this study, we took advantage of the large repertoire of genome sequences that is available for the genus *Bradyrhizobium* to reveal that the RhcI-T3SS gene clusters and the *nod* genes largely share a common evolutionary history and have both been conserved in almost all nodulating *Bradyrhizobium* strains. This concomitant maintenance of *nod* and T3SS clusters is most likely driven by host plant-related selective pressure and suggests a much more important role of the T3SS in the symbiotic process than hitherto suspected. The recent identification that a T3SS effector (ErnA) that is conserved among bradyrhizobia is able to induce nodule formation in the absence of NFs strengthens this idea [25].

This comparative genomic study provides new opportunities to fill gaps in our sparse knowledge of the role of the T3SS in rhizobial symbiosis. Indeed, we propose new genetic determinants that are probably required for the T3SS symbiotic function based on their phylogenetic distribution: *nopE* and *nopH* distribution complements that of *nopX*, in association with their respective RhcI gene cluster types, suggesting a possible role for these two proteins that is related to the translocation function. Furthermore, we highlighted the high diversity of NopM and SUMO protease effector families, revealing a much larger effectome than anticipated in some strains. Some of the newly identified putative effectors may hence possibly contribute to symbiotic efficiency and/

or the modulation of the host spectrum of these strains. The discovery of several atypical T3SS clusters in some strains raises the question of whether these gene clusters contribute to the symbiotic properties of bradyrhizobia. Finally, the realization that some strains isolated from nodules lack *nod* genes but contain a T3SS cluster raises the possibility that the T3SS evolved as the main determinant governing the symbiotic interaction. These strains might constitute new models for future studies aiming at a better understanding of T3SS's symbiotic function.

Funding information
This work was supported by grants from the French national research agency (grant 'SymEffectors' number ANR-16-CE20-0013). A. T. was supported by a PhD fellowship from the French Ministry of National Education, Higher Education and Research. F. L. was supported by a UK Medical Research Council (MRC) grant (MR/N010760/1).

Acknowledgements
The LABGeM (CEA/Genoscope and CNRS UMR8030), the France Génomique and French Bioinformatics Institute national infrastructures (funded as part of the Investissement d'Avenir programme managed by Agence Nationale pour la Recherche, contracts ANR-10-INBS-09 and ANR-11-INBS-0013) are acknowledged for support with the MicroScope annotation platform. Computational calculations were performed on the Imperial College high-performance computing (HPC) cluster and on the MRC Cloud Infrastructure for Microbial Bioinformatics (MRC CLIMB) cloud-based computing servers.

Author contributions
A.T and E.G. devised the study, with input from F.L., Z.R., D.G., A.C and R. K. Bioinformatic analyses were performed by A.T and F.L. with advice from R.K., A.T., F.L and E.G. wrote the manuscript, and all authors provided critical review and commentary.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ *et al*. A global atlas of the dominant bacteria found in soil. *Science* 2018;359:320–325.
2. VanInsberghe D, Maas KR, Cardenas E, Strachan CR, Hallam SJ *et al*. Non-symbiotic *Bradyrhizobium* ecotypes dominate North American forest soils. *ISME J* 2015;9:2435–2441.
3. Jones FP, Clark IM, King R, Shaw LJ, Woodward MJ *et al*. Novel European free-living, non-diazotrophic *Bradyrhizobium* isolates from contrasting soils that lack nodulation and nitrogen fixation genes - a genome comparison. *Sci Rep* 2016;6:25858.
4. Oldroyd GED, Speak OGE. Speak, Friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat Rev Microbiol* 2013;11:252–263.
5. Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E *et al*. Legumes symbioses: absence of NOD genes in photosynthetic bradyrhizobia. *Science* 2007;316:1307–1312.
6. Masson-Boivin C, Giraud E, Perret X, Batut J. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol* 2009;17:458–466.
7. Masson-Boivin C, Sachs JL. Symbiotic nitrogen fixation by rhizobia-the roots of a success story. *Curr Opin Plant Biol* 2018;44:7–15.
8. Zhao R, Liu LX, Zhang YZ, Jiao J, Cui WJ *et al*. Adaptive evolution of rhizobial symbiotic compatibility mediated by co-evolved insertion sequences. *ISME J* 2018;12:101–111.
9. Kaneko T, Nakamura Y, Sato S, Minamisawa K, Uchiumi T *et al*. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res* 2002;9:189–197.
10. Reeve W, van Berkum P, Ardley J, Tian R, Gollagher M *et al*. High-quality permanen^{t}draf^{t} genome sequence of the *Bradyrhizobium elkanii* type strain USDA 76^{T}, isolated from *Glycine max* (L.) Merr. *Stand Genomic Sci* 2017;12:26.
11. Gully D, Teulet A, Busset N, Nouwen N, Fardoux J *et al*. 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* 2018;5:e00421–17.
12. Le Quéré A, Gully D, Teulet A, Navarro E, Gargani D *et al*. Complete genome sequence of *Bradyrhizobium* sp. strain ORS3257, an efficient nitrogen-fixing bacterium isolated from cowpea in Senegal. *Microbiol Resour Announc* 2019;8:e01449–18.
13. Deng W, Marshall NC, Rowland JL, McCoy JM, Worrall LJ *et al*. Assembly, structure, function and regulation of type III secretion systems. *Nat Rev Microbiol* 2017;15:323–337.
14. Büttner D. Behind the lines-actions of bacterial type III effector proteins in plant cells. *FEMS Microbiol Rev* 2016;40:894–937.
15. Deakin WJ, Broughton WJ. Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat Rev Microbiol* 2009;7:312–320.
16. Staehelin C, Krishnan HB. Nodulation outer proteins: double-edged swords of symbiotic rhizobia. *Biochem J* 2015;470:263–274.
17. Tampakaki AP. Commonalities and differences of T3SSs in rhizobia and plant pathogenic bacteria. *Front Plant Sci* 2014;5:114.
18. Krause A, Doerfel A, Göttfert M. Mutational and transcriptional analysis of the type III secretion system of *Bradyrhizobium japonicum*. *Mol Plant Microbe Interact* 2002;15:1228–1235.
19. Wassem R, Kobayashi H, Kambara K, Le Quéré A, Walker GC *et al*. TtsI regulates symbiotic genes in *Rhizobium* species NGR234 by binding to *tts* boxes. *Mol Microbiol* 2008;68:736–748.
20. Miwa H, Okazaki S. How effectors promote beneficial interactions. *Curr Opin Plant Biol* 2017;38:148–154.
21. Yang S, Tang F, Gao M, Krishnan HB, Zhu H. R gene-controlled host specificity in the legume-rhizobia symbiosis. *Proc Natl Acad Sci USA* 2010;107:18735–18740.
22. Sugawara M, Takahashi S, Umehara Y, Iwano H, Tsurumaru H *et al*. Variation in bradyrhizobial NopP effector determines symbiotic incompatibility with Rj2-soybeans via effector-triggered immunity. *Nat Commun* 2018;9:3139.
23. Okazaki S, Kaneko T, Sato S, Saeki K. Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *Proc Natl Acad Sci USA* 2013;110:17131–17136.
24. Okazaki S, Tittabutr P, Teulet A, Thouin J, Fardoux J *et al*. Rhizobium-legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. *ISME J* 2016;10:64–74.
25. Teulet A, Busset N, Fardoux J, Gully D, Chaintreuil C *et al*. The rhizobial type III effector eRNA confers the ability to form nodules in legumes. *Proc Natl Acad Sci USA* 2019;116:21758–21768.
26. Lassalle F, Veber P, Jauneikaite E, Didelot X. Automated reconstruction of all gene histories in large bacterial pangenome datasets and search for co-evolved gene modules with Pantagruel. *bioRxiv* 2019;586495.
27. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 2017;35:1026–1028.
28. Sievers F, Higgins DG, Desmond GH. Clustal omega for making accurate alignments of many protein sequences. *Protein Sci* 2018;27:135–145.
29. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* 2006;34:W609–W612.
30. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Phylogenetics* 2013;30:1312–1313.

31. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A *et al*. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 2012;61:539–542.
32. Vallenet D, Calteau A, Cruveiller S, Gachet M, Lajus A *et al*. Microscope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res* 2017;45:D517–D528.
33. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018;35:1547–1549.
34. Cohen O, Ashkenazy H, Belinky F, Huchon D, Pupko T. GLOOME: gain loss mapping engine. *Bioinformatics* 2010;26:2914–2915.
35. Szöllősi GJ, Tannier E, Daubin V, Boussau B. The inference of gene trees with species trees. *Syst Biol* 2015;64:e42–e62.
36. Morel B, Kozlov AM, Stamatakis A, Szöllősi GJ. GeneRax: a tool for species tree-aware maximum likelihood based gene tree inference under gene duplication, transfer, and loss. *bioRxiv* 2019;779066.
37. Connor TR, Loman NJ, Thompson S, Smith A, Southgate J *et al*. CLIMB (the cloud infrastructure for microbial bioinformatics): an online resource for the medical microbiology community. *Microb Genom* 2016;2:e000086.
38. Okazaki S, Noisangiam R, Okubo T, Kaneko T, Oshima K *et al*. Genome analysis of a novel *Bradyrhizobium* sp. DOA9 carrying a symbiotic plasmid. *PLoS One* 2015;10:e0117392.
39. Iida T, Itakura M, Anda M, Sugawara M, Isawa T *et al*. Symbiosis island shuffling with abundant insertion sequences in the genomes of extra-slow-growing strains of soybean bradyrhizobia. *Appl Environ Microbiol* 2015;81:4143–4154.
40. Miché L, Moulin L, Chaintreuil C, Contreras-Jimenez JL, Munive-Hernández J-A *et al*. 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* 2010;12:2152–2164.
41. Okubo T, Fukushima S, Itakura M, Oshima K, Longtonglang A *et al*. Genome analysis suggests that the soil oligotrophic bacterium *Agromonas oligotrophica* (*Bradyrhizobium oligotrophicum*) is a nitrogen-fixing symbiont of *Aeschynomene indica*. *Appl Environ Microbiol* 2013;79:2542–2551.
42. Tian R, Parker M, Seshadri R, Reddy T, Markowitz V *et al*. High-quality permanent draft genome sequence of *Bradyrhizobium* sp. Tv2a.2, a microsymbiont of *Tachigali versicolor* discovered in Barro Colorado Island of Panama. *Stand Genomic Sci* 2015;10:27.
43. Stępkowski T, Watkin E, McInnes A, Gurda D, Gracz J *et al*. Distinct *Bradyrhizobium* [corrected] communities nodulate legumes native to temperate and tropical monsoon Australia. *Mol Phylogenet Evol* 2012;63:265–277.
44. van Berkum P, Fuhrmann JJ. Evolutionary relationships among the soybean bradyrhizobia reconstructed from 16S rRNA gene and internally transcribed spacer region sequence divergence. *Int J Syst Evol Microbiol* 2000;50 Pt 6:2165–2172.
45. Benzine J, Shelobolina E, Xiong MY, Kennedy DW, McKinley JP *et al*. Fe-phyllosilicate redox cycling organisms from a redox transition zone in Hanford 300 area sediments. *Front Microbiol* 2013;4:388.
46. Crovadore J, Calmin G, Cochard B, Chablais R, Lefort F. Whole-genome sequence of *Bradyrhizobium elkanii* strain UASWS1015, a highly ammonia-tolerant nitrifying bacterium. *Genome Announc* 2016;4:e00111–00116.
47. van Berkum P, Elia P, Song Q, Eardly BD. Development and application of a multilocus sequence analysis method for the identification of genotypes within genus *Bradyrhizobium* and for establishing nodule occupancy of soybean (*Glycine max* L. Merr). *Mol Plant Microbe Interact* 2012;25:321–330.
48. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR *et al*. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461–466.
49. de Lajudie PM, Andrews M, Ardley J, Eardly B, Jumas-Bilak E *et al*. Minimal standards for the description of new genera and species of rhizobia and agrobacteria. *Int J Syst Evol Microbiol* 2019;69:1852–1863.
50. Avontuur JR, Palmer M, Beukes CW, Chan WY, Coetzee MPA *et al*. Genome-informed *Bradyrhizobium* taxonomy: where to from here? *Syst Appl Microbiol* 2019;42:427–439.
51. Garrido-Sanz D, Redondo-Nieto M, Mongiardini E, Blanco-Romero E, Durán D *et al*. Phylogenomic analyses of *Bradyrhizobium* reveal uneven distribution of the lateral and subpolar flagellar systems, which extends to *Rhizobiales*. *Microorganisms* 2019;7:50.
52. Troisfontaines P, Cornelis GR. Type III secretion: more systems than you think. *Physiology* 2005;20:326–339.
53. Hu Y, Huang H, Cheng X, Shu X, White AP *et al*. A global survey of bacterial type III secretion systems and their effectors. *Environ Microbiol* 2017;19:3879–3895.
54. Schmeisser C, Liesegang H, Krysciak D, Bakkou N, Le Quéré A *et al*. *Rhizobium* sp. strain NGR234 possesses a remarkable number of secretion systems. *Appl Environ Microbiol* 2009;75:4035–4045.
55. Weidner S, Becker A, Bonilla I, Jaenicke S, Lloret J *et al*. Genome sequence of the soybean symbiont *Sinorhizobium fredii* HH103. *J Bacteriol* 2012;194:1617–1618.
56. Schuldes J, Rodriguez Orbegoso M, Schmeisser C, Krishnan HB, Daniel R *et al*. Complete genome sequence of the broad-host-range strain *Sinorhizobium fredii* USDA257. *J Bacteriol* 2012;194:4483.
57. Gazi AD, Sarris PF, Fadouloglou VE, Charova SN, Mathioudakis N *et al*. Phylogenetic analysis of a gene cluster encoding an additional, rhizobial-like type III secretion system that is narrowly distributed among *Pseudomonas syringae* strains. *BMC Microbiol* 2012;12:188.
58. Hueck CJ. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* 1998;62:379–433.
59. Okazaki S, Zehner S, Hempel J, Lang K, Göttfert M. Genetic organization and functional analysis of the type III secretion system of *Bradyrhizobium elkanii*. *FEMS Microbiol Lett* 2009;295:88–95.
60. Kimbrel JA, Thomas WJ, Jiang Y, Creason AL, Thireault CA *et al*. Mutualistic co-evolution of type III effector genes in *Sinorhizobium fredii* and *Bradyrhizobium japonicum*. *PLoS Pathog* 2013;9:e1003204.
61. Hempel J, Zehner S, Göttfert M, Patschkowski T. Analysis of the secretome of the soybean symbiont *Bradyrhizobium japonicum*. *J Biotechnol* 2009;140:51–58.
62. Jiménez-Guerrero I, Pérez-Montaño F, Medina C, Ollero FJ, López-Baena FJ. NopC is a rhizobium-specific type 3 secretion system effector secreted by *Sinorhizobium* (*Ensifer*) *fredii* HH103. *PLoS One* 2015;10:e0142866.
63. Büttner D, Nennstiel D, Klüsener B, Bonas U. Functional analysis of HrpF, a putative type III translocon protein from *Xanthomonas campestris* pv. *vesicatoria*. *J Bacteriol* 2002;184:2389–2398.
64. Marie C, Deakin WJ, Ojanen-Reuhs T, Diallo E, Reuhs B *et al*. TtsI, a key regulator of *Rhizobium* species NGR234 is required for type III-dependent protein secretion and synthesis of rhamnose-rich polysaccharides. *Mol Plant Microbe Interact* 2004;17:958–966.
65. Helene LCF, Delamuta JRM, Ribeiro RA, Hungria M. *Bradyrhizobium mercantei* sp. nov., a nitrogen-fixing symbiont isolated from nodules of *Deguelia costata* (syn. Lonchocarpus costatus). *Int J Syst Evol Microbiol* 2017;67:1827–1834.
66. Ribeiro RA, Helene LCF, Delamuta JRM, Hungria M. Genome sequence of *Bradyrhizobium mercantei* strain SEMIA6399^T, isolated from nodules of *Deguelia costata* in Brazil. *Genome Announc* 2017;5:e00943–17.
67. Ji H, Dong H. Key steps in type III secretion system (T3SS) towards translocon assembly with potential sensor at plant plasma membrane. *Mol Plant Pathol* 2015;16:762–773.
68. Saad MM, Staehelin C, Broughton WJ, Deakin WJ. Protein-protein interactions within type III secretion system-dependent pili of *Rhizobium* sp. strain NGR234. *J Bacteriol* 2008;190:750–754.
69. Pesce C, Jacobs JM, Berthelot E, Perret M, Vancheva T *et al*. Comparative genomics identifies a novel conserved protein, HpaT, in proteobacterial type III secretion systems that do not

possess the putative translocon protein HrpF. *Front Microbiol* 2017;8:1177.

70. Wenzel M, Friedrich L, Göttfert M, Zehner S. The type III-secreted protein NopE1 affects symbiosis and exhibits a calcium-dependent autocleavage activity. *Mol Plant Microbe Interact* 2010;23:124–129.
71. Büttner D, Bonas U. Port of entry--the type III secretion translocon. *Trends Microbiol* 2002;10:186–192.
72. Kambara K, Ardissone S, Kobayashi H, Saad MM, Schumpp O *et al*. Rhizobia utilize pathogen-like effector proteins during symbiosis. *Mol Microbiol* 2009;71:92–106.
73. Dean P. Functional domains and motifs of bacterial type III effector proteins and their roles in infection. *FEMS Microbiol Rev* 2011;35:1100–1125.
74. Gomez-Valero L, Rusniok C, Carson D, Mondino S, Pérez-Cobas AE *et al*. More than 18 000 effectors in the *Legionella* genus genome provide multiple, independent combinations for replication in human cells. *Proc Natl Acad Sci USA* 2019;116:2265–2273.
75. Xin D-W, Liao S, Xie Z-P, Hann DR, Steinle L *et al*. Functional analysis of NopM, a novel E3 ubiquitin ligase (Nel) domain effector of Rhizobium sp. strain NGR234. *PLoS Pathog* 2012;8:e1002707.
76. Xu C-C, Zhang D, Hann DR, Xie Z-P, Staehelin C. Biochemical properties and in planta effects of NopM, a rhizobial E3 ubiquitin ligase. *J Biol Chem* 2018;293:15304–15315.
77. Rodrigues JA, López-Baena FJ, Ollero FJ, Vinardell JM, Espuny MDR *et al*. NopM and NopD are rhizobial nodulation outer proteins: identification using LC-MALDI and LC-ESI with a monolithic capillary column. *J Proteome Res* 2007;6:1029–1037.
78. Faruque OM, Miwa H, Yasuda M, Fujii Y, Kaneko T *et al*. Identification of *Bradyrhizobium elkanii* genes involved in incompatibility with soybean plants carrying the *Rj4* allele. *Appl Environ Microbiol* 2015;81:6710–6717.
79. Dai W-J, Zeng Y, Xie Z-P, Staehelin C. Symbiosis-promoting and deleterious effects of NopT, a novel type 3 effector of *Rhizobium* sp. strain NGR234. *J Bacteriol* 2008;190:5101–5110.
80. Fotiadis CT, Dimou M, Georgakopoulos DG, Katinakis P, Tampakaki AP. Functional characterization of NopT1 and NopT2, two type III effectors of *Bradyrhizobium japonicum*. *FEMS Microbiol Lett* 2012;327:66–77.
81. Hoyer E, Knöppel J, Liebmann M, Steppert M, Raiwa M *et al*. Calcium binding to a disordered domain of a type III-secreted protein from a coral pathogen promotes secondary structure formation and catalytic activity. *Sci Rep* 2019;9:7115.

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.

Article : “Identification of type III effectors modulating the symbiotic properties of *Bradyrhizobium vignae* strain ORS3257 with various *Vigna* species”

scientific reports

Check for updates

OPEN

Identification of type III effectors modulating the symbiotic properties of *Bradyrhizobium vignae* strain ORS3257 with various *Vigna* species

Pongpan Songwattana[1,5], Clémence Chaintreuil[2,3,5], Jenjira Wongdee[1], Albin Teulet[2], Mamadou Mbaye[2], Pongdet Piromyou[1], Djamel Gully[2], Joel Fardoux[2], Alexandre Mahougnon Aurel Zoumman[3,4], Alicia Camuel[2], Panlada Tittabutr[1], Neung Teaumroong[1]✉ & Eric Giraud[2]✉

The *Bradyrhizobium vignae* strain ORS3257 is an elite strain recommended for cowpea inoculation in Senegal. This strain was recently shown to establish symbioses on some *Aeschynomene* species using a cocktail of Type III effectors (T3Es) secreted by the T3SS machinery. In this study, using a collection of mutants in different T3Es genes, we sought to identify the effectors that modulate the symbiotic properties of ORS3257 in three *Vigna* species (*V. unguiculata*, *V. radiata* and *V. mungo*). While the T3SS had a positive impact on the symbiotic efficiency of the strain in *V. unguiculata* and *V. mungo*, it blocked symbiosis with *V. radiata*. The combination of effectors promoting nodulation in *V. unguiculata* and *V. mungo* differed, in both cases, NopT and NopAB were involved, suggesting they are key determinants for nodulation, and to a lesser extent, NopM1 and NopP1, which are additionally required for optimal symbiosis with *V. mungo*. In contrast, only one effector, NopP2, was identified as the cause of the incompatibility between ORS3257 and *V. radiata*. The identification of key effectors which promote symbiotic efficiency or render the interaction incompatible is important for the development of inoculation strategies to improve the growth of *Vigna* species cultivated in Africa and Asia.

Thanks to the ability of legume plants to interact symbiotically with nitrogen fixing rhizobia, they play a major agronomical and ecological role. This interaction results in the formation of symbiotic organs, called nodules, generally on the root system, in which the bacteria fix nitrogen for the plant's benefit and, in exchange, receive carbon sources.

The establishment of a successful symbiosis relies on a complex signaling dialogue between the two partners thereby enabling their mutual recognition, the triggering of the programmes that lead to the formation of the nodule and its infection, and the weakening of the plant immune system allowing the plant to tolerate a large intracellular bacterial population inside the nodule cells[1].

The Nod factors (NF) synthetized by the rhizobia in response to the perception of plant flavonoid signal, are recognised as the first "key" signal governing nodulation and infection processes[2]. Beyond the NF signals, other bacterial components are important for the successful establishment of the symbiosis, including exopolysaccharides (EPS), lipopolysaccharides (LPS), capsular polysaccharides (KPS), and cyclic β-glucans[3,4]. These surface

[1]School of Biotechnology, Institute of Agricultural Technology, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand. [2]IRD, Laboratoire des Symbioses Tropicales et Méditerranéennes, UMR 113, IRD/CIRAD/INRAE/Université de Montpellier/SupAgro, Campus de Baillarguet, TA-A82/J, 34398 Montpellier Cedex 5, France. [3]IRD, Laboratoire Commun de Microbiologie, UR040, ISRA, UCAD, Centre de Recherche de Bel Air, Dakar, Senegal. [4]Département de Biologie Végétale, University Cheikh Anta Diop, Dakar, Senegal. [5]These authors contributed equally: Pongpan Songwattana and Clémence Chaintreuil. ✉email: neung@sut.ac.th; eric.giraud@ird.fr

polysaccharides can act directly as a symbiotic signal or may be modified to avoid the induction of plant triggered immunity (PTI) which normally occurs when the plant recognises a PAMP signal[5,6].

More surprisingly, to cope with the plant immune system, similar to plant and animal pathogens, rhizobia use T3SS to interact with their host[7]. The T3SS is a complex needle-like machine that allows injection of specialised bacterial effector proteins directly into the host cell[8]. These effectors modulate key host cell functions, most of which are related to the suppression of host immunity but others can modulate plant hormone signaling, metabolism or organelle function[9].

Not all rhizobia possess a T3SS, but surprisingly, this secretory machinery is widespread among rhizobia belonging to the *Bradyrhizobium* genus[10]. Indeed, a phylogenomic study, based on a large number of bradyrhizobial genome sequences currently available showed that more than 90% of the nodulating bradyrhizobia strains contained not only *nod* genes but also a T3SS[10]. In fact, it appears that the *nod* and T3SS genes share the same evolutionary history, bradyrhizobia acquired the two gene clusters simultaneously via horizontal transfer of a symbiotic island containing both. The systematic conservation of the T3SS in almost all nodulating *Bradyrhizobium* strains, suggests that, like *nod* genes, T3SS genes play an important role in the symbiotic process.

The rhizobial type III effector proteins (T3Es), including those called Nop for "Nodulation outer proteins", can play a positive, negative or neutral role in the symbiotic process depending on the host plant and the rhizobium[11,12]. They can promote bacterial infection and nodulation by repressing specific defense responses of plants activated by the recognition of bacterial signals (PAMP), or conversely, they can make the interaction incompatible if they are recognized by immune receptors of the plants [resistance proteins (R)] leading to activation of effector-triggered immunity (ETI)[13,14].

More recently, it has been shown that some legume species of the *Aeschynomene* genus and also the cultivar *Glycine max* cv. Enrie are nodulated by *Bradyrhizobium* strains even if NF synthesis is disrupted[15,16]. In this case, the establishment of the interaction requires that the bacteria has a functional T3SS, evidence that certain T3Es can also directly activate nodulation in some legumes by bypassing the NF signal.

To better understand which effectors are important for the establishment of this NF-independent, T3SS-dependent symbiosis, we almost exhaustively mutated the predicted effector gene repertoire (23 out of the 27 putative T3Es) in *Bradyrhizobium* sp. strain ORS3257, which is one of the most efficient strains in triggering the NF-independent symbiotic process in *A. indica*[17]. Analysis of the symbiotic properties of the mutants demonstrated that in this strain, T3SS-dependent symbiosis relies on a cocktail of at least five effectors which play synergistic and complementary roles in nodule organogenesis, and in the infection and repression of plant immune responses. Among them, we identified the nuclear-targeted ErnA effector, which is highly conserved among bradyrhizobia, as a key actor of nodule organogenesis[17].

Interestingly, the strain ORS3257, which was originally isolated from a root nodule of *Vigna unguiculata*, is one of the two dominant genetic types of *Bradyrhizobium* strains nodulating cowpea at different sites in Senegal[18,19]. Considering that this genome has 98.07% average nucleotide identity (ANI) with that of *B. vignae* LMG28791 type strain[20] (calculated using http://enve-omics.ce.gatech.edu/ani/), this strain belongs to the *B. vignae* species. Furthermore, the strain was shown to be one of the most promising inoculants because it induces efficient nitrogen fixing nodules on various *V. unguiculata* cultivars and it is a good competitor for nodule occupancy[18,21]. However, nothing is known about the importance of *nod,* T3SS and T3Es genes in the symbiotic properties of this strain with its natural host plant.

In this study, we took advantage of a collection of mutants of ORS3257 strain to identify which genes are relevant for the symbiotic efficiency of this strain with *V. unguiculata*. Furthermore, because the strain has the potential to be used as inoculant, we also examined the symbiotic properties of the mutants on two other *Vigna* species, *V. radiata* (mung bean) and *V. mungo* (urad bean), which are cultivated in Asia.

Results

The T3SS of *B. vignae* ORS3257 strain has a different impact on symbiosis depending on the *Vigna* species considered. In the first approach, we used three *Vigna* species to compare the symbiotic properties of the ORS3257 WT strain and its T3SS derivative mutant (3257Δ*T3SS*) obtained by deletion of the main *rhc* gene cluster (from *nopB* to *rhcU*) (Supplementary Fig. S1). Interestingly, in the three plant species tested, we observed a strong effect of the T3SS deletion but this effect differed with the species. As shown in Fig. 1a,b,d,e,g,h,j,k, the *V. unguiculata* and *V. mungo* plants inoculated with the 3257Δ*T3SS* mutant had lower chlorophyll content, lower growth, as well as fewer nodules than plants inoculated with the WT strain. These observations show that T3SS plays a positive role in the symbiotic efficiency of the strain in these two plant species.

In contrast, while the WT strain could not induce nodules on the *V. radiata* species, 3257Δ*T3SS* formed nodules that were functional in terms of benefits for plant growth and leaf chlorophyll content (Fig. 1c,f,i,l).

To go deeper into the characterisation of the effect of the T3SS mutation, we also performed cytological analysis of the nodules elicited by the mutant. As shown in Fig. 1m–x, the nodules elicited by the 3257Δ*T3SS* mutant displayed no apparent disorders whatever the species tested, i.e. the nodules were the characteristic pink color of leghemoglobin, evidence for their ability to fix nitrogen. Furthermore, confocal observations showed that the nodule cells of the central tissue were perfectly infected intracellularly by viable bacteria as revealed by the green colour of the Cyto9 stain (Fig. 1y–ah). The only notable observation was the presence of yellow autofluorescence zones in some rare nodules of *V. mungo,* suggesting the accumulation of phenolic compounds due to plant defence reactions (Fig. 1ab).

Taken together these data indicate that the T3SS contributes to the symbiotic efficiency of *B. vignae* on *V. mungo* and *V. unguiculata* while it renders incompatible the symbiotic interaction with *V. radiata* most probably by plant recognition of a specific T3E inducing ETI.

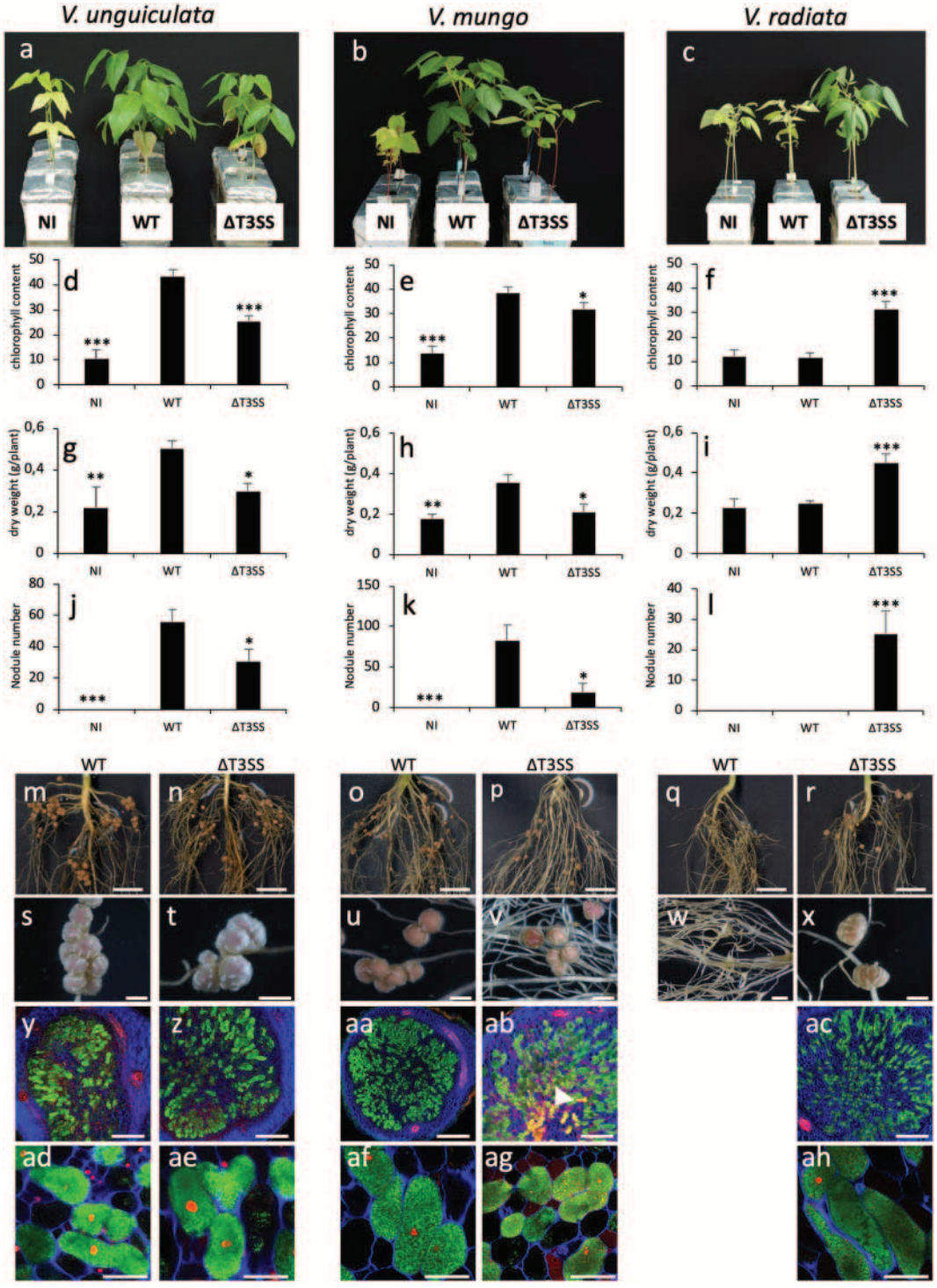


Figure 1. Symbiotic properties of *Bradyrhizobium vignae* ORS3257 strain (WT) and its derivative T3SS mutant (ΔT3SS) in different *Vigna* species. (**a**–**c**) Comparison of the growth of the plants (aerial part), non-inoculated (NI) or inoculated with ORS3257 wild-type strain (WT) or its derivative T3SS mutant (Δ*T3SS*) at 21 days after inoculation; (**d**–**f**) Leaf chlorophyll content; (**g**–**i**) Plant dry weight; (**j**–**l**) Number of nodules per plant. The experiment was carried out in duplicate with five plants per condition. *P < 0.05, **P < 0.01, and ***P < 0.001, significant differences between WT ORS3257 and the T3SS mutant using a nonparametric Kruskal–Wallis test. (**m**–**x**) View of the root and the nodules induced by ORS3257 strain and its derivative T3SS mutant (Δ*T3SS*). (Scale bars, m to r, 1 cm; s to x, 2 mm). (**y**–**ah**) Cytological analysis of the nodules induced by strain WT ORS3257 and Δ*T3SS* observed by confocal microscopy after staining with SYTO 9 (green: live bacteria), calcofluor (blue: plant cell wall), and propidium iodide (red; infected plant nuclei and dead bacteria or bacteria with compromised membranes). (Scale bars, y to ac, 200 µm; ad to ah, 50 µm). White arrowhead in AB shows the necrotic zone.

Considering that the T3SS has been shown to bypass the requirement of NF signalling during the interactions of some *Bradyrhizobium* strains including ORS3257 with some tropical legumes[15,16], we constructed a mutant in which the *nodABC* genes were deleted and analysed its symbiotic properties. The ability of the 3257Δ*nodABC* mutant to form nodules on *V. mungo* and *V. unguiculata* was found to be completely aborted. Like the WT strain, it did not also induce nodules on *V. radiata*. These data indicate that the symbiotic process between ORS3257 and *Vigna* species (at least for *V. mungo* and *V. unguiculata*) depends on NFs.

Effectors NopT and NopAB play a major positive role in the interaction between ORS3257 and *V. mungo* and *V. unguiculata*. To identify which T3Es have a positive impact on the symbiotic relationship between ORS3257 and *V. unguiculata* and *V. mungo*, we analysed the symbiotic properties of the T3Es ORS3257 mutants we had previously constructed[17]. This consisted of a collection of ten mutants, five deletion mutants in which genomic regions containing several clustered T3E genes were deleted by double crossing over, and five insertional mutants in specific T3E genes (*nopBW*, *nopL*, *nopP2*, *ernA*, and *Brad7238*) obtained by insertion of the non-replicative pVO155 vector by single crossing over (Supplementary Fig. S1). These ten mutants include 23 out of the 27 effectors predicted in silico in the ORS3257 genome[17]. Out of the 10 tested mutants, two mutants (3257Δ*regB* and 3257Δ*regD*) were drastically impacted in the two *Vigna* species, while two additional mutants (3257Δ*regA*, and 3257Ω*nopP2*) displayed a weak phenotype only in *V. mungo* (Fig. 2a–f).

Given the symbiotic defects observed in the 3257Δ*regB* and 3257Δ*regD* mutants that were even greater than those observed in the Δ*T3SS* mutant, we analysed sub-mutants of these two regions to identify the effector(s) responsible for the phenotype.

Region B contains two putative effector genes (*nopT* and *nopAO*) as well as two additional *tts* boxes with no clearly defined downstream coding sequence (Fig. 3a). Analysis of four available sub-mutants of this region corresponding to two insertional mutants in *nopT* and *nopAO* (Ω*nopT* and Ω*nopAO*) and two deletion mutations (Δ*regB1* and Δ*regB2*) encompassing the region surrounding the two *tts* boxes showed that only the Ω*nopT* mutant had a symbiotic defect in the two *Vigna* species (Fig. 3b–e and Supplementary Fig. S2b–e). Nevertheless, it revealed that the phenotype of the Ω*nopT* mutant was less drastic than the one of the Δ*regB* mutant. In particular, while the Δ*regB* mutant induced only a few white nodules in the two *Vigna* species, the Ω*nopT* mutant formed more nodules and these nodules consisted of a mixture of white and pink nodules in *V. unguiculata* and mainly pink nodules in *V. mungo* (Fig. 3f–k and Supplementary Fig. S2f–k). Furthermore, cytological analysis revealed that the Δ*regB* nodules were completely necrotic and/or very poorly infected, with a notable problem of releasing of bacteria into the host cells, i.e. the bacteria were not uniformly dispersed in the cell but remained clustered together (Fig. 3l–o and Supplementary Fig. S2l–o). While the white Ω*nopT* nodules on *V. unguiculata* or the pink Ω*nopT* nodules on *V. mungo* were generally better infected, some plant cells were filled with uniformly dispersed viable bacteria (Fig. 3p,q and Supplementary Fig. S2p,q). These data suggest that the absence of *nopT* in the Δ*regB* mutant was largely responsible for the observed Δ*regB* phenotype but the absence of another as yet unidentified genetic determinant in this region likely also contributes to the observed defects.

The region deleted in the 3257Δ*regD* mutant corresponds to a small region (3.5 kb) containing two putative effector genes (*nopAB* and *Brad3257_7707*) (Fig. 4a). The analyses of the two available insertional mutants in these effectors showed that the mutation in *Brad3257_7707* did not lead to any particular phenotype whereas the symbiotic properties of the Ω*nopAB* mutant were strongly affected in the two *Vigna* species tested (Fig. 4b–k and Supplementary Fig. S3b–k). However, as observed for the Δ*regB* and Ω*nopT* mutants, the absence of the *nopAB* gene alone cannot explain the phenotype of the Δ*regD* mutant. Indeed, the Δ*regD* mutant produced fewer nodules than the Ω*nopAB* mutant (Fig. 4e and Supplementary Fig. S3e). Furthermore, the Δ*regD* nodules in both *Vigna* species displayed the same irregular infection as that observed in Δ*regB* nodules, while such a defect of infection appeared less marked in the Ω*nopAB* nodules, in particular in *V. mungo* (Fig. 4l–q and Supplementary Fig. S3l–q). It is therefore possible that the more drastic phenotype of the Δ*regD* mutant was due to a cumulative effect of deleting *nopAB* and *Brad3257_7707* genes.

We also analysed the symbiotic phenotypes of two sub-mutants (Ω*nopP*1 and Ω*nopM1)* of region A that are specifically mutated in one of the two effectors identified in this region (Supplementary Fig. S4a). While the two mutants showed no particular phenotypes when tested on *V. unguiculata,* when tested on *V. mungo* both showed a significant deficiency in some of the symbiotic parameters measured (Supplementary Fig. S4b–g). This strongly suggests that the weak phenotype observed in the Δ*regA* mutant in *V. mungo* resulted from the combined absence of NopP1 and NopM1.

Taken together, these data indicate that both NopT and NopAB play a major positive role in the symbiotic interaction between ORS3257 and the two *Vigna* species *V. unguiculata* and *V. mungo*. But beyond these two T3Es, we hypothesise that other effectors such as NopP1, NopM1, Brad3257_7707 and very probably others that remain to be identified, also play a role in the global symbiotic performance of the strain, in particular in *V. mungo*.

NopP2 blocks the symbiotic interaction between ORS3257 and *V. radiata* cv. SUT1. As done previously, we tested the 10 constructed mutants on *V. radiata* cv. SUT1 to identify the effector(s) which compromise the symbiosis of ORS3257 with this *Vigna* species. All the mutants tested displayed the same strict *nod* minus phenotype as observed in the WT strain except for the Ω*nopP2* mutant (Fig. 2g–i). Like the Δ*T3SS* mutant, the *nopP2* mutant resulted in functional pink nodules as indicated by improved plant growth and increased leaf chlorophyll content (Fig. 5a,d,g,j,o,p). These data clearly show that the effector NopP2 is responsible for the symbiotic incompatibility between ORS3257 and *V. radiatia* cv. SUT1.

Interestingly, NopP was also recently identified in *B. diazoefficiens* USDA122 as the causal effector that prevents the symbiotic interaction between this strain and *Rj2*-Soybeans[14]. Surprisingly, by swapping the *nopP*

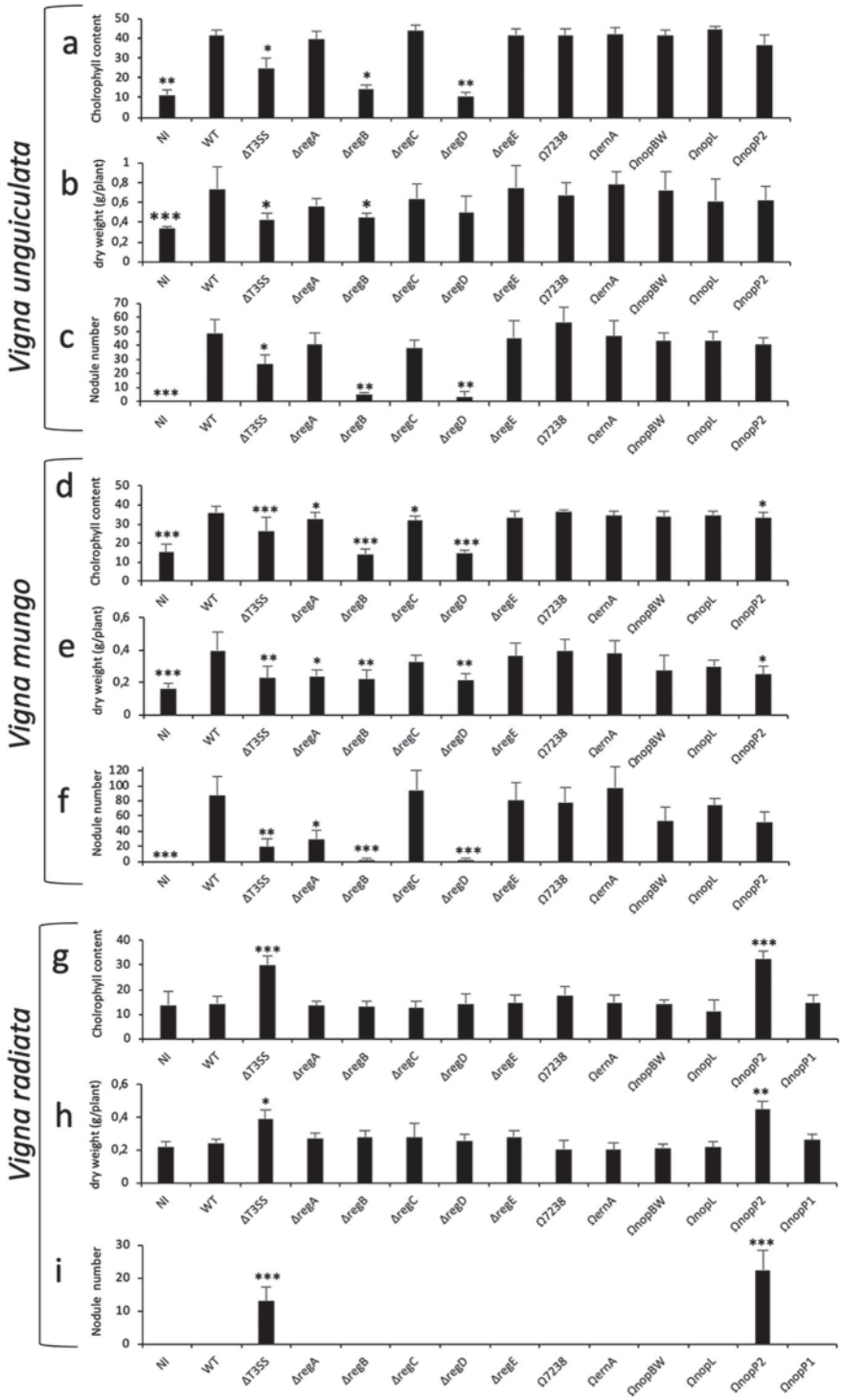


Figure 2. Identification of T3Es in *Bradyrhizobium vignae* ORS3257 strain that modulate the symbiotic interaction with *V. unguiculata, V. mungo* and *V. radiata*. (**a**,**d**,**g**) Leaf chlorophyll content of the plants inoculated with ORS3257 wild-type strain (WT) and its mutant derivatives at 21 days after inoculation; (**b**,**e**,**h**) Plant dry weight; (**c**,**f**,**i**) Number of nodules per plant. The experiment was carried out in duplicate with five plants per condition. *P < 0.05, **P < 0.01, and ***P < 0.001, significant differences between WT ORS3257 and the mutants using a nonparametric Kruskal–Wallis test.

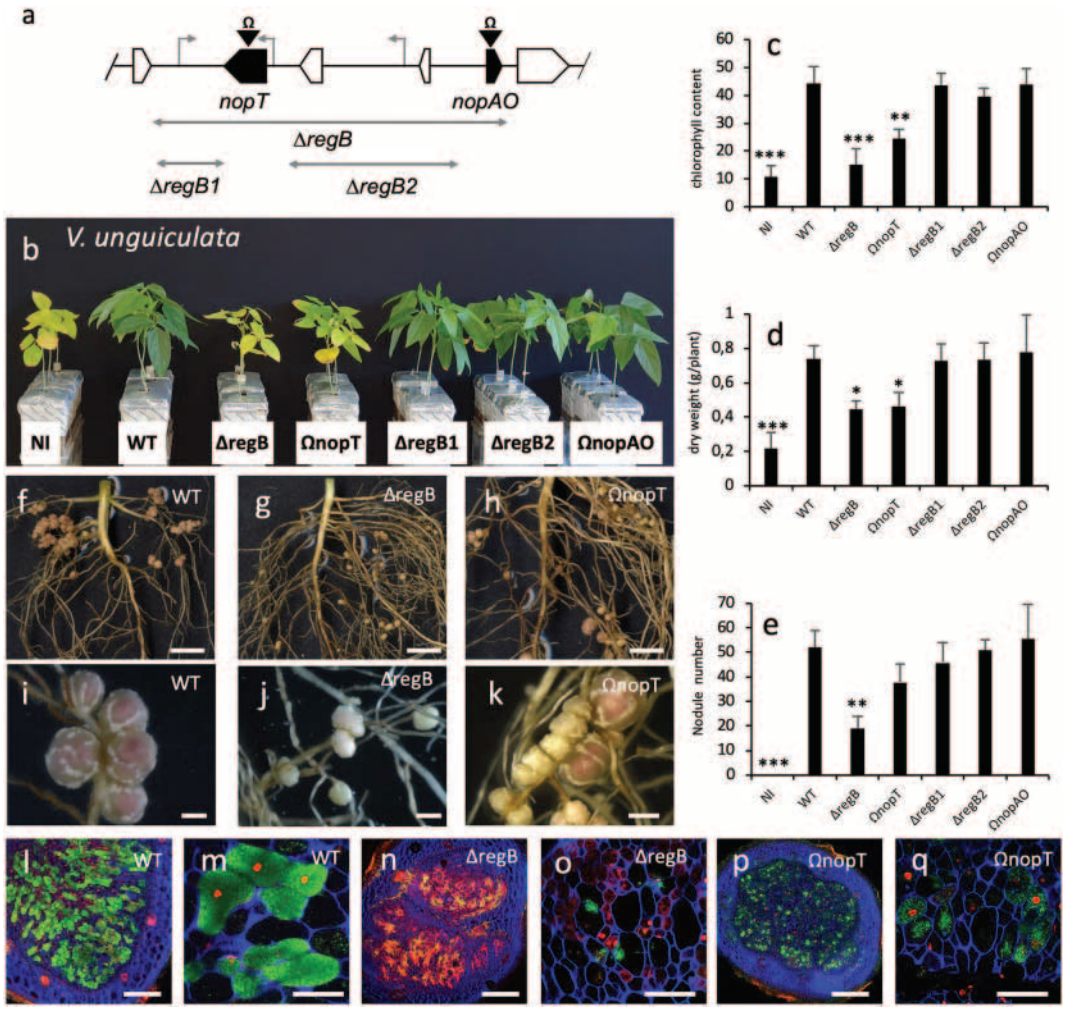


Figure 3. Symbiotic properties on *V. unguiculata* of the ORS3257 mutants in the different effectors and the regions encompassing the *tts* boxes identified in region B. (**a**) Genetic organisation of the putative effectors and *tts* boxes identified in region B. The deleted regions in the mutants are indicated by double direction arrows. The insertion mutants are indicated by black arrowheads carrying the Ω sign. In black, putative effector genes; gray arrows, *tts* boxes. (**b**) Comparison of the growth of the plants (aerial part), non-inoculated (NI) or inoculated with ORS3257 wild-type strain (WT) or its derivative mutants at 21 days after inoculation. (**c**) Leaf chlorophyll content. (**d**) Plant dry weight; (**e**) Number of nodules per plant. The experiment was carried out in duplicate with five plants per condition. *P < 0.05, **P < 0.01, and ***P < 0.001, significant differences between WT ORS3257 and its derivative mutants using a nonparametric Kruskal–Wallis test. (**f**–**k**) View of the root and the nodules induced by strain ORS3257 and its derivative mutants. (Scale bars: f to h, 1 cm; i to k, 2 mm). (**l**–**q**) Cytological analysis of the nodules induced by strain ORS3257 and its derivative mutants observed by confocal microscopy after staining with SYTO 9 (green: live bacteria), calcofluor (blue: plant cell wall), and propidium iodide (red: infected plant nuclei and dead bacteria or bacteria with compromised membranes). (Scale bars, l, n, p, 200 µm; m, o, q, 50 µm).

gene in USDA122 with the one of the compatible *B. diazoefficiens* USDA110 strain in which the corresponding protein differs in only four amino acid residues, it has been possible to restore a symbiotic interaction, indicating that the plant can reject or accept a compatible symbiont simply by monitoring some minor variations in NopP[14]. Interestingly, the ORS3257 strain contains two *nopP* homologues, *Brad3257_6978 (nopP1)* in region A, and *Brad3257_7831* (*nopP2*) elsewhere on the symbiotic island (Supplementary Fig. S1). The *ΔregA* mutant was unable to form nodules, indicating that NopP1 is not involved in limiting this symbiosis (Fig. 2i), which we confirmed by also testing a specific insertional mutant (*ΩnopP1*) that also displayed a strict *nod* minus phenotype (Fig. 2i). Sequence alignment between the different NopP sequences showed that NopP2 shared higher identity with the NopP from USDA122 and USDA110 (71.9% of identity with each) than with NopP1 (40.6% and 39.9% of identity with the NopP of USDA122 and USDA110, respectively) (Supplementary Fig. S5). However, two out of the three AA of USDA122 NopP shown to be required for *Rj2*-soybean incompatibility are not conserved in NopP2 (Supplementary Fig. S5).

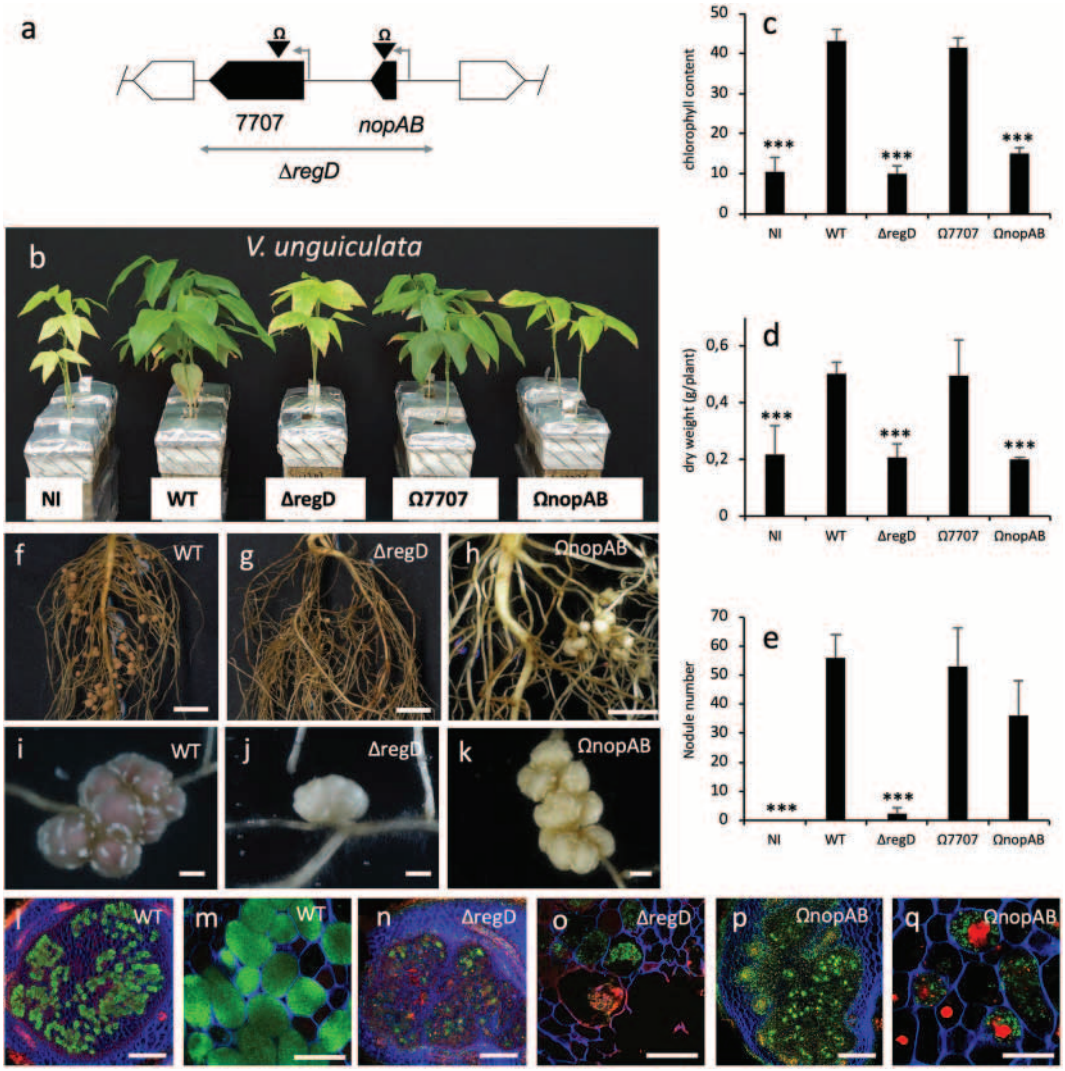


Figure 4. Symbiotic properties of the ORS3257 mutants in the different effectors of the region D in *V. unguiculata*. (**a**) Genetic organisation of the putative effectors and *tts* boxes identified in region D. The deleted region in the mutant is indicated by a double direction arrow. The insertion mutants are indicated by black arrowheads carrying the Ω sign. In black, putative effector genes; gray arrows, *tts* boxes. (**b**) Comparison of the growth of the plants (aerial part), non-inoculated (NI) or inoculated with ORS3257 wild-type strain (WT) or its derivative mutants at 21 days after inoculation. (**c**) Leaf chlorophyll content. (**d**) Plant dry weight; (**e**) Number of nodules per plant. The experiment was carried out in duplicate with five plants per condition. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$, significant differences between WT ORS3257 and its derivative mutants using a nonparametric Kruskal–Wallis test. (**f**–**k**) View of the root and the nodules induced by strain ORS3257 and its derivative mutants. (Scale bars: f to h, 1 cm; i to k, 2 mm.) (**l**–**q**) Cytological analysis of the nodules induced by strain ORS3257 and its derivative mutants observed by confocal microscopy after staining with SYTO 9 (green: live bacteria), calcofluor (blue: plant cell wall), and propidium iodide (red: infected plant nuclei and dead bacteria or bacteria with compromised membranes). (Scale bars, l, n, p, 200 µm; m, o, q, 50 µm).

To investigate whether the mechanisms governing the incompatibility between ORS3257 and *V. radiata* cv. SUT1 are the same as those described between USD122 and *Rj2* soybeans, we first analysed the symbiotic properties of USDA110 and USDA122 strains in *V. radiata* cv. SUT1. USDA110 was clearly seen to be more efficient than USDA122 in this cultivar, i.e. more nodules were formed, and plant growth and chlorophyll content were higher (Fig. 5a,d,g,j). But at the same time, this difference was not as black and white as the difference observed between ORS3257 and Ω*nopP2*. Indeed, the nodules elicited by USDA110 and USDA122 look similar, both were pink and we even observed a small improvement in growth and chlorophyll content in the plants inoculated with USDA122 (Fig. 5d,g,m,n).

The ORS3257 WT strain and its Δ*T3SS* and Ω*nopP2* mutants were also tested on *Rj2* and *rj2* soybeans, and USDA110 and USDA122 strains were also included as controls. We confirmed the inability of USDA122 to

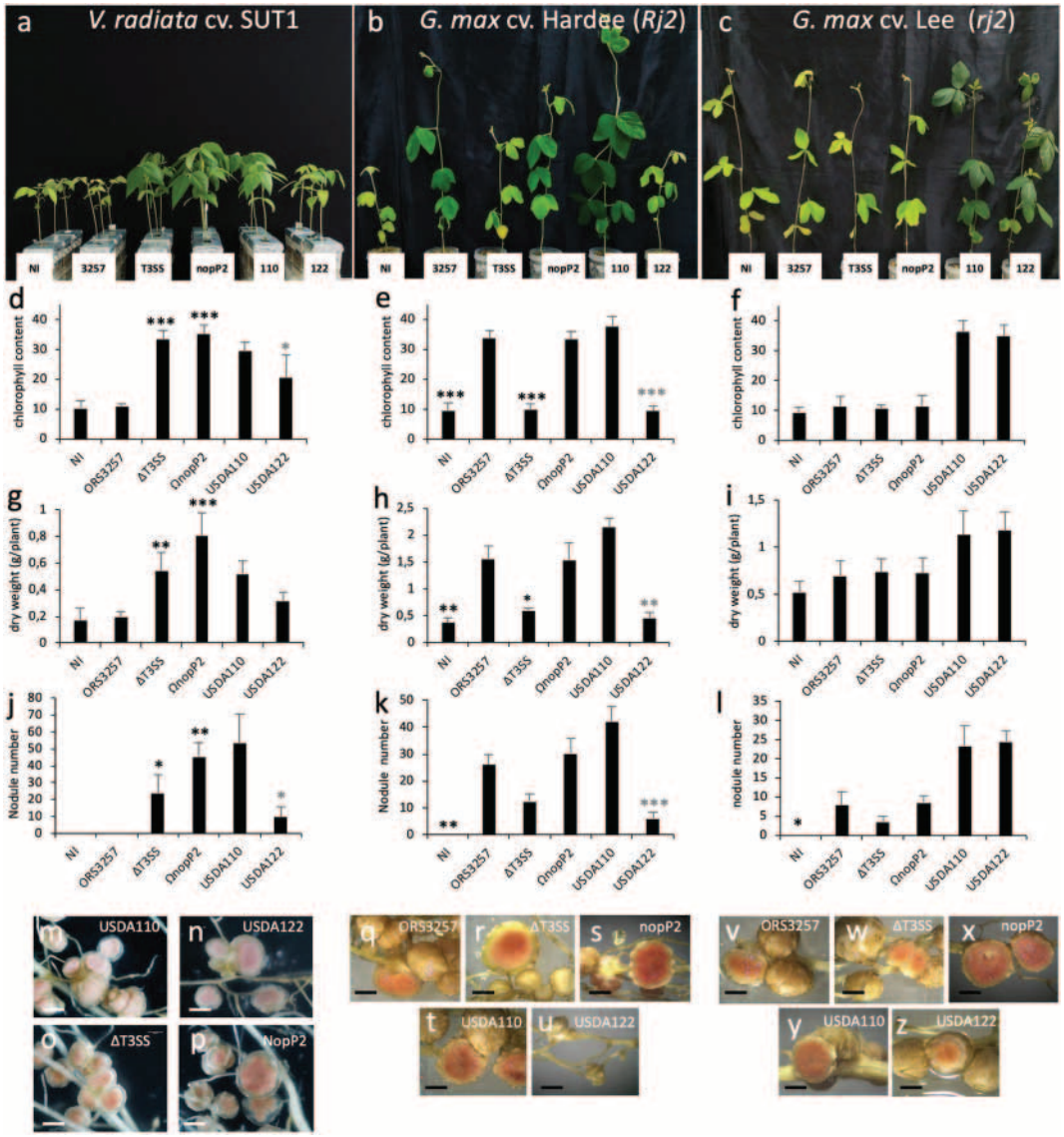


Figure 5. Symbiotic properties of various *Bradyrhizobium* strains in *V. radiata* cv. SUT1, *G. max* cv. Hardee (*Rj2*) and *G. max* cv. Lee (*rj2*). (**a**–**c**) Comparison of the growth of the plants (aerial part), non-inoculated (NI) or inoculated with ORS3257 wild-type strain (3257) or its derivative mutants (ΔT3SS and ΩNopP2) or with USDA122 (122) or USDA110 (110) WT-strains at 21 days after inoculation. (**d**–**f**) Leaf chlorophyll content; (**g**–**i**) Plant dry weight; (**j**–**l**) Number of nodules per plant. The experiment was carried out in duplicate with five plants per condition. *P < 0.05, **P < 0.01, and ***P < 0.001. Black asterisks indicate significant differences between WT ORS3257 and other strains tested, using a nonparametric Kruskal–Wallis test, Gray asterisks indicate significant differences between WT USDA110 and WT USDA122. (**m**–**z**) View of the nodules. (Scale bars, 1 mm).

efficiently nodulate *Rj2*-soybean 'Hardee', only a few small pseudo-nodules were observed and no improvement in plant fitness was measured (Fig. 5b,e,h,k,t,u), whereas when tested on *rj2*-soybean 'Lee', both strains displayed similar symbiotic efficiency (Fig. 5c,f,i,l,y,z). Unexpectedly, the ORS3257 WT strain and its Ω*nopP2* mutant were found to be symbiotically efficient in *Rj2*-soybean 'Hardee' while the Δ*T3SS* mutant induced fewer nodules that had no apparent benefit for plant growth and chlorophyll content (Fig. 5b,e,h,k,q,r,s). These data indicate that NopP2 protein of ORS3257 was not recognized as an incompatible effector by *Rj2*-soybean 'Hardee' and that by secreting other T3Es, T3SS played an important positive role in this symbiotic interaction. T3SS was also observed to play a weak positive role in the nodulation of *rj2*-soybean 'Lee', as ORS3257 and the Ω*nopP2* mutant resulted in more nodules than the Δ*T3SS* mutant (Fig. 5l). Nevertheless, in the latter case, the nodules elicited by WT ORS3257 and Ω*nopP2* were apparently not functional since no improvement was observed in plant growth or chlorophyll content despite the fact that the nodules displayed a pink colour (Fig. 5c,f,i,v,w,x).

Taken together, these data clearly show that NopP2 of ORS3257 blocks nodulation with *V. radiata* cv. SUT1 but it is still not clear whether this symbiotic incompatibility is induced via a R-protein homologous to Rj2 identified in soybeans.

Discussion

A recent phylogenomic analysis conducted on a large number of *Bradyrhizobium* genomes showed that a T3SS is almost systematically found in all the strains nodulating and containing *nod* genes, suggesting that T3SS is an overriding factor in determining the symbiotic properties of the rhizobia that belong to this genus[10]. Our study conducted on *B. vignae* ORS3257 strain supports this hypothesis as we observed drastic impacts of the mutation of the T3SS machinery on the symbiosis of the bacteria on the three *Vigna* species tested. Interestingly, and as already reported for several *Bradyrhizobium* or rhizobium strains belonging to other genera[11], depending on the host plant, T3SS can either support symbiotic efficiency, or quite the reverse, completely block the symbiosis.

By using a collection of mutants affected in the effector genes identified in this strain, we showed that a combination of different numbers of effectors is required to support the efficiency of the symbiosis of the ORS3257 strain in *V. unguiculata* and *V. mungo*, while only one effector NopP2 is sufficient to block the symbiosis with *V. radiata* cv. SUT1. Interestingly, NopT and NopAB, which were found to play an important role in the symbiosis of ORS3257 with *V. unguiculata* and *V. mungo*, as well as NopP1 and NopM1, which were found to play a positive role in *V. mungo* although to a lesser extent, were previously shown to be required for the establishment of the NF-independent symbiosis with *Aeschynomene indica*[17]. Conversely, ErnA, which was the key effector triggering nodulation on *A. indica*, did not play a significant role in these two *Vigna* species. The NFs signal synthetized by ORS3257 strain, which we showed to be absolutely indispensable for the symbiosis in the two latter species, most probably renders the role of ErnA in nodule organogenesis superfluous. However, from these observations, we are still not able to rule that ErnA is dispensable in all the symbioses involving NFs, given the very high occurrence of this gene in *Bradyrhizobium* strains containing *nod* genes and its high level of conservation[10,17]. It is possible that ErnA strengthens or replaces the NF signal in certain culture conditions or during interactions with particular hosts.

The NopT and NopAB mutants induced fewer nodules and some disorders in intracellular infection were observed, particularly in *V. unguiculata*. In *A. indica*, it was previously reported that the nodules formed by these two mutants were completely devoid of infected central tissue, thereby pointing to an important role for these effectors in bacterial infection[17]. The importance of NopT in the modulation of the symbiotic interaction has also already been observed in *Enisifer fredii* NGR234. Interestingly, in this strain, NopT has been shown to act as a 'two-edge sword' which affected nodulation either positively in *Phaseolus vulgaris* or negatively in *Crotalaria juncea*[22]. The NopT effectors identified in rhizobia belong to the YopT-AvrPphB cysteine protease family initially characterised in pathogenic bacteria. Several functional characterisation of the NopT from diverse rhizobia (*E. fredii* NGR234, *Mesorhizobium amphore* and *B. diazoefficiens*) show that this T3E displays autoproteolytic activity, targets the plasma membrane of the host cell most probably through its lipid acylation after its translocation, and elicits HR-like cell death when expressed in tobacco[22–24]. Very recently, two possible targets of the NopT of *M. amphore* CCNWGS0123 were identified in *Robinia pseudoacacia*, an ATP-citrate synthase and a hypersensitive induced response protein (HIRP)[24]. However, how the interaction of NopT with these targets can modulate the symbiotic process remains to be clarified.

Knowledge of the NopAB effector is still extremely limited. This effector family was originally identified in several *B. diazoefficiens* strains through an in silico search of genome sequences of genes associated with the regulatory *tts*-box motif[25]. Sequence analysis of NopAB did not enable prediction of any functional domains and no homolog has been found outside of the bradyrhizobia, suggesting that NopAB is a specific effector of unknown function that emerged in this genus. Interestingly, an analysis of its distribution in the *Bradyrhizobium* genus showed that NopAB is very well conserved since more than 50% of the strains containing a T3SS have a NopAB homolog (47 out of 92 strains)[10]. However, to our knowledge, no other study has reported a role for this effector in other *Bradyrhizobium* strains. Considering the important role played by NopAB in ORS3257 during the interaction with several legume species, it would be interesting to further explore its function and examine its symbiotic role in other strains. To a lesser extent, we also observed that NopM1 and NopP1 play a significant role during the symbiotic interaction of ORS3257 and *V. mungo*. These two effectors were also shown to play a positive role during the interaction between *E. fredii* NGR234 and several legumes species[26,27], while in *B. elkanii* USDA61 it has been recently reported that NopM induces a nodule early senescence-like response in *Lotus burttii* and *L. japonicus* MG-20[28]. The mode of action of NopP remains unclear. It has been shown that the NopP of NGR234 can be phosphorylated in vitro by plant kinases but it is still not known whether this phosphorylation suppresses plant defence responses by interfering with MAP kinase signalling pathways as has been shown for NopL, another effector in NGR234, which also plays a positive symbiotic role in certain host plants[29,30]. The NopM effector of NGR234 has recently been shown also to be phosphorylated by plant kinase[31]. Interestingly, this effector displays moreover an E3 ubiquitin ligase activity[31,32]. It has been hypothesised that NopM promotes nodulation by ubiquitination of specific host proteins to target them for proteasome-dependent proteolysis and/or by interfering with the MAP signalling pathways, like NopL[31]. Similarly, the NopM1 and NopP1 identified in ORS3257 could interfere via phosphorylation and ubiquitination and act synergistically to suppress plant defences and subsequently favours nodulation of *V. mungo*. It is important to note that ORS3257 possesses another homolog of NopP (NopP2) and NopM (NopM2). The Ω*nopP2* mutant was found to be slightly affected during the symbiosis with *V. mungo* (Fig. 2d,e), in combination with NopP1 and NopM1, NopP2 could therefore also be involved in the modulation of plant immunity. In the case of NopM2, at this stage, in the absence of a specific mutant in this gene, it is difficult to argue whether or not this effector has a cumulative effect, even if the *ΔregC* mutant, which includes the deletion of *nopM2*, displayed no specific phenotype (see discussion below).

Several observations lead us to assume that during the screening of this collection of mutants, we missed certain effectors perhaps involved in the modulation of the symbiotic properties of ORS3257. One observation is the fact that the phenotype of the Δ*T3SS* mutant, in which the secretory machinery is completely abolished, was not as drastic as in other mutants in some deleted regions (Δ*regB* and Δ*regD*) or in specific effectors (Ω*nopT* and Ω*nopAB*). In the Δ*T3SS* mutant, all the possible effectors that could have either a positive or a negative effect were eliminated. We therefore hypothesise that the less drastic phenotype observed results from the suppression of effectors playing a negative role. However, none of the mutants tested showed a symbiotic gain. It cannot be ruled out that the mutagenesis approach we used, which consisted in deleting regions containing several *T3Es*, has certain limits if effectors in this same region have antagonistic effects on the symbiosis. In particular, this could be the case of region C that contains nine T3Es and for which the corresponding mutant Δ*regC* displayed no particular phenotype. In the same line of thought, we observed that the Δ*regB* and Δ*regD* mutants had a more drastic phenotype than the Ω*nopT* and Ω*nopAB* mutants, suggesting that the absence of other effectors or some genetic determinants present in these regions also contribute to the phenotype of the two deleted mutants. But at the same time, the mutation of the other candidate effectors (*Brad3257_7707*, *nopAO)*, or the deletion of the regions encompassing the two *tts* boxes (*ΔregB1* and *ΔregB2*) did not lead to a specific phenotype. It is possible that some effectors or determinants in these two regions make an incremental contribution with NopT and NopAB in the global symbiotic performance of the strain but that their specific effect remains limited.

Surprisingly, the effectors identified as playing an important role during the symbiotic interaction between ORS3257 and *V. mungo* are not the same than the ones recently identified in *B. elkanii* USDA61 interacting with the same species[33]. In particular, T3SS was shown to be essential in USDA61 during this interaction and the T3E shown to play a major positive role is NopL, while the Ω*nopL* mutant of ORS3257 displayed no specific phenotype in this plant species. Similarly, NopAB, which is important in ORS3257, is absent in USDA61. For NopT, the question remains whether this effector plays a similar role in USDA61 for which a homolog has been found but, to our knowledge, no mutant has yet been tested in *V. mungo*.

On the other hand, USDA61 was also shown to be incompatible on *V. radiata* cv. KPS1, but the causal effector that blocks this interaction is not NopP but another effector named InnB for (Incompatible nodulation B)[34,35]. It is to note that USDA61 possess two NopP homologs, one of which is very similar to NopP2 of ORS3257 (Supplementary Fig. S5). We cannot exclude the possibility that these NopP homologs have a negative effect during this interaction but if it exists, the effect is very likely limited given that *innB* mutant is highly efficient[35]. The observed differences between the two strains could be explained by the fact that the plant cultivar used in the two studies was different, but also by the fact that each strain has evolved and adapted its effectome according to the different plant partners encountered during its history.

It has been proposed that the molecular mechanisms mediated by recognition of NopP by Rj2 protein leading to the incompatibility between some *Bradyrhizobium* strains and some soybean cultivars are conserved in other legumes[14]. Two arguments were put forward to support this hypothesis: (i) the identification of Rj2 orthologs in various legume genomes including in *V. radiata* and, (ii) the fact that a USDA122 mutant in which the *nopP* gene was swapped by the one of USDA110 was found to be more efficient in *V. radiata* cv KPS1 than the WT strain USDA122. The observation that NopP2 in ORS3257 blocks the interaction with another *V. radiata* ecotype is consistent with this view. Nevertheless, we observed that NopP2 of ORS3257 did not restrict the compatibility with the Rj2 soybean 'Hardee'. It is possible that the Rj2 orthologs identified in *V. radiata* and soybeans evolved differently and recognize different specific NopP variations in order to monitor the compatibility or incompatibility of their symbionts. It would be interesting to investigate if ORS3257 is able to nodulate other *V. radiata* varieties and compare the Rj2 sequences between compatible and incompatible ecotypes.

While inoculation of soybean fields is a common practice to increase grain yields across the world, inoculation of mung bean, cowpea or black mung bean fields remains limited, especially in Africa[36]. Beside the low cost of chemical fertilizers which represents a real impediment to the development of sustainable agriculture, one limitation in the use of inoculant is the identification of effective strain that are more competitive than the endogenous microflora and can be used for a broader range of crops. Identifying the genetic determinants in Rhizobium which promote or compromise the symbiosis is important for the design or selection of appropriate inoculum strains.

Methods

Bacterial strains and culture conditions. Supplementary Table S1 lists the bacterial mutants of the ORS3257 strain used in this study. *Bradyrhizobium* sp. strains (ORS3257, USDA110 and USDA122) and the derivative mutants of ORS3257 were grown in yeast mannitol (YM) medium[37] at 34 °C on a rotary shaker at 150 rpm. When required, the media were supplemented with appropriate antibiotics at the following concentrations: 20 µg/mL Nalidixic acid, 50 µg/mL kanamycin (Km), and 20 µg/mL cefotaxime (Cefo). The cefotaxime resistance of the mutants results to the use of the cefotaxime resistant gene originating from *E. coli* strain 25,104 as selective marker[16].

Construction of a nodABC mutant of ORS3257. For the construction of the 3257Δ*nodABC* mutant, the upstream and downstream flanking regions (around 750 bp) of the *nodABC* genes were amplified using the primers GGGGGGGATCCAATAGCAAACATCAGTTTGGAAAAG (Up.NodA.3257.f), GAGCCAATCAAAGCT TGCCGGGTTTCCTTGCGCCAGTGGAAT (Up.NodA.3257.r), GAAACCCGGCAAGCTTTGATTGGCTCA TGATTCCGGACGCAA (Dw.NodC.3257.f), and CGTTGCTCTAGATTTCTTGCTCGATGCAAAAGACAAG (Dw.NodC.3257.r). The PCR fragments were merged by overlap extension PCR and cloned into pNPTS129 at the BamHI/XbaI sites. Subsequently, a cefotaxime resistance cartridge was cloned between the upstream and downstream fragments in the HindIII site. The resulting plasmid was then transferred by conjugation into the ORS3257 strain and the deleted mutant obtained by double crossing over was selected as described previously[38].

Plant nodulation and symbiosis analysis. The plants used in this study are listed in Supplementary Table S2. The seeds of the different legume species tested were surface sterilised with a 3% (v/v) calcium hypochlorite solution for 3 min, washed with sterilised water and then soaked in 70% (v/v) ethanol for 3 additional min. Finally, the seeds were washed in sterilised water five times and left to germinate on 0.8% water agar plates at 34 °C overnight. The germinated seeds were transplanted into Leonard jars filled with sterilised vermiculite[39] and watered with BNM medium[40]. The plantlets were grown under the following controlled environmental conditions: 28 °C with a 16-h light and 8-h dark regime at light intensities of 300 µmol/m^{2}S and 70% humidity. Five days after planting, each seedling was inoculated with 1 mL of a 5-day old inoculum after washing and adjusting the optical density at 600 nm to 1 (approximately 10^9 cells/mL). Twenty-one days after inoculation (dpi), the leaf chlorophyll content was monitored using a Minolta SPAD-502 chlorophyll meter[41], the number of nodules was counted, and the dry weights of plants were determined after drying at 50 °C for 96 h. The experiment was carried out in duplicate using five plants per condition.

Cytological analysis of nodules. Fresh nodules were observed under an optical Macroscope (Nikon AZ100, Champigny-sur-Marne, France). Sections (40–50 µm thick) of fresh nodules were prepared using a VT1200S vibratome (Leica Nanterre, France). Nodule sections were incubated for 15 min in live/dead staining solution (5 µM SYTO 9 and 30 µM propidium iodide in 50 mM Tris pH 7.0 buffer; Live/Dead BacLight, Invitrogen, Carlsbad, CA, USA). The sections were then removed and incubated for an additional 15 min in 10 mM phosphate saline buffer containing calcofluor white M2R (Sigma, Munich, Germany) at a final concentration of 0.01% (wt/vol) to stain the plant cell wall[42]. Analyses were carried out using a confocal laser-scanning microscope (Carl Zeiss LSM 700, Jena, Germany). Calcofluor was excited at 405 nm with emission signal collection at 405–470 nm. For SYTO 9 and propidium iodide, an excitation wavelength of 488 and 555 nm was used to collect emission signals at 490–522 nm and 555–700 nm, respectively. Images were obtained using the ZEN 2008 software (Zeiss, Oberkochen, Germany).

Received: 8 December 2020; Accepted: 5 February 2021
Published online: 01 March 2021

References

1. Poole, P., Ramachandran, V. & Terpolilli, J. Rhizobia: From saprophytes to endosymbionts. *Nat. Rev. Microbiol.* **16**, 291–303 (2018).
2. Oldroyd, G. E. Speak, friend, and enter: Signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* **11**, 252–263 (2013).
3. Gibson, K., Kobayashi, H. & Walker, G. C. Molecular determinants of a symbiotic chronic infection. *Annu. Rev. Genet.* **42**, 413–441 (2008).
4. Janczarek, M., Rachwał, K., Marzec, A., Grzadziel, J. & Palusinska-Szysz, M. Signal molecules and cell-surface components involved in early stages of the legume-rhizobium interactions. *Appl. Soil Ecol.* **85**, 94–113 (2015).
5. Kawaharada, Y. *et al.* Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature* **523**, 308–312 (2015).
6. Gourion, B., Berrabah, F., Ratet, P. & Stacey, G. Rhizobium-legume symbioses: The crucial role of plant immunity. *Trends Plant. Sci.* **20**, 186–194 (2015).
7. Deakin, W. J. & Broughton, W. J. Symbiotic use of pathogenic strategies: Rhizobial protein secretion systems. *Nat. Rev. Microbiol.* **7**, 312–320 (2009).
8. Deng, W. *et al.* Assembly, structure, function and regulation of type III secretion systems. *Nat. Rev. Microbiol.* **15**, 323–337 (2017).
9. Macho, A. P. Subversion of plant cellular functions by bacterial type-III effectors: Beyond suppression of immunity. *New Phytol.* **210**, 51–57 (2016).
10. Teulet, A. *et al.* Phylogenetic distribution and evolutionary dynamics of *nod* and T3SS genes in the genus *Bradyrhizobium*. *Microb. Genom.* https://doi.org/10.1099/mgen.0.000407 (2020).
11. Staehelin, C. & Krishnan, H. B. Nodulation outer proteins: Double-edged swords of symbiotic rhizobia. *Biochem. J.* **470**, 263–274 (2015).
12. Miwa, H. & Okazaki, S. How effectors promote beneficial interactions. *Curr. Opin. Plant Biol.* **38**, 148–154 (2017).
13. Yang, S., Tang, F., Gao, M., Krishnan, H. B. & Zhu, H. R gene-controlled host specificity in the legume-rhizobia symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 18735–18740 (2010).
14. Sugawara, M. *et al.* Variation in bradyrhizobial NopP effector determines symbiotic incompatibility with Rj2-soybeans via effector-triggered immunity. *Nat. Commun.* **9**, 3139. https://doi.org/10.1038/s41467-018-05663-x (2018).
15. Okazaki, S., Kaneko, T., Sato, S. & Saeki, K. Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *Proc. Natl. Acad. Sci. USA* **110**, 17131–17136 (2013).
16. Okazaki, S. *et al.* Rhizobium-legume symbiosis in the absence of Nod factors: Two possible scenarios with or without the T3SS. *ISME J.* **10**, 64–74 (2016).
17. Teulet, A. *et al.* The rhizobial type III effector ErnA confers the ability to form nodules in legumes. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21758–21768 (2019).
18. Krasova-Wade, T. *et al.* Diversity of indigeneous bradyrhizobia associated with 3 cowpea cultivars (*Vigna unguiculata* (L.) Walp.) grow under limited and favorable water conditions in Senegal (West Africa). *Afr. J. Biotechnol.* **21**, 13–22 (2003).
19. Krasova-Wade, T. *et al.* Eco-geographical diversity of cowpea bradyrhizobia in Senegal is marked by dominance of two genetic types. *Syst. Appl. Microbiol.* **37**, 129–139 (2014).
20. Grönemeyer, J. L., Hurek, T., Bünger, W. & Reinhold-Hurek, B. *Bradyrhizobium vignae* sp. nov., a nitrogen-fixing symbiont isolated from effective nodules of Vigna and Arachis. *Int. J. Syst. Evol. Microbiol.* **66**, 62–69 (2016).
21. Krasova-Wade, T. *et al.* Water-condition effects on rhizobia competition for cowpea nodule occupancy. *Afr. J. Biotechnol.* **5**, 1457–1463 (2006).
22. Dai, W. J., Zeng, Y., Xie, Z. P. & Staehelin, C. Symbiosis promoting and deleterious effects of NopT, a novel type 3 effector of *Rhizobium* sp. NGR234. *J. Bacteriol.* **190**, 5101–5110 (2008).
23. Fotiadis, C. T., Dimou, M., Georgakopoulos, D. G., Katinakis, P. & Tampakaki, A. P. Functional characterization of NopT1 and NopT2, two type III effectors of *Bradyrhizobium japonicum*. *FEMS Microbiol. Lett.* **327**, 66–77 (2011).

24. Luo, Y. *et al.* Identification of *Robinia pseudoacacia* target proteins responsive to *Mesorhizobium amphore* CCNWGS0123 effector protein NopT. *J. Exp Bot.* https://doi.org/10.1093/jxb/eraa405 (2020).
25. Kimbrel, J. A. *et al.* Mutualistic co-evolution of type III effector genes in *Sinorhizobium fredii* and *Bradyrhizobium japonicum*. *PLoS Pathog.* **9**, e1003204. https://doi.org/10.1371/journal.ppat.1003204 (2013).
26. Marie, C. *et al.* Characterization of Nops, nodulation outer proteins, secreted via the type III secretion system of NGR234. *Mol. Plant Microbe Interact.* **16**, 743–751 (2003).
27. Kambara, K. *et al.* Rhizobia utilize pathogen-like effector proteins during symbiosis. *Mol. Microbiol.* **71**, 92–106 (2009).
28. Kusakabe, S. *et al. Lotus* accessions possess multiple checkpoints triggered by different type III secretion system effectors of the wide-host-range symbiont *Bradyrhizobium elkanii* USDA61. *Microbes Environ.* **35**, 1–16 (2020).
29. Skorpil, P. *et al.* NopP, a phosphorylated effector of *Rhizobium* sp. strain NGR234, is a major determinant of nodulation of the tropical legumes *Flemingia congesta* and *Tephrosia vogelii*. *Mol. Microbiol.* **57**, 1304–1317 (2005).
30. Zhang, L., Chen, X. J., Lu, H. B., Xie, Z. P. & Staehelin, C. Functional analysis of the type 3 effector nodulation outer protein L (NopL) from *Rhizobium* sp. NGR234: Symbiotic effects, phosphorylation, and interference with mitogen-activated protein kinase signaling. *J. Biol. Chem.* **286**, 32178–32187 (2011).
31. Xu, C. C., Zhang, D., Hann, D. R., Xie, Z. P. & Staehelin, C. J. Biochemical properties and in planta effects of NopM, a rhizobial E3 ubiquitin ligase. *Biol. Chem.* **293**, 15304–15315 (2018).
32. Xin, D. W. *et al.* Functional analysis of NopM, a novel E3 ubiquitin ligase (NEL) domain effector of *Rhizobium* sp. strain NGR234. *PLoS Pathog.* **8**, e1002707. https://doi.org/10.1371/journal.ppat.1002707 (2012).
33. Nguyen, H.P., Ratu, S.T.N., Yasuda, M., Teaumroong, N. & Okazaki, S. Identification of *Bradyrhizobium elkanii* USDA61 Type III effectors determining symbiosis with *Vigna mungo*. *Genes (Basel)*. **11**, 474. https://doi.org/10.3390/genes11050474.PMID:32349348 (2020).
34. Nguyen, H.P., Miwa, H., Kaneko, T., Sato, S. & Okazaki, S. Identification of *Bradyrhizobium elkanii* genes involved in incompatibility with *Vigna radiata*. *Genes (Basel)*. **8**, 374. https://doi.org/10.3390/genes8120374.PMID:29292795 (2017).
35. Nguyen, H. P., Ratu, S. T. N., Yasuda, M., Göttfert, M. & Okazaki, S. InnB, a Novel Type III Effector of *Bradyrhizobium elkanii* USDA61 controls symbiosis with *Vigna* Species. *Front. Microbiol.* **9**, 3155. https://doi.org/10.3389/fmicb.2018.03155 (2018).
36. Santos, M. S., Nogueira, M. A. & Hungria, M. Microbial inoculants: Reviewing the past, discussing the present and previewing an outstanding future for the use of beneficial bacteria in agriculture. *AMB Express.* **9**, 205. https://doi.org/10.1186/s13568-019-0932-0 (2019).
37. Vincent, J. *A Manual for the Practical Study of Root-Nodule Bacteria* (Blackwell Scientific Publications, Oxford, 1970).
38. Giraud, E., Lavergne, J. & Verméglio, A. Characterization of bacteriophytochromes from photosynthetic bacteria: Histidine kinase signaling triggered by light and redox sensing. *Methods Enzymol.* **471**, 135–159 (2010).
39. Somasegaran, P. & Hoben, H. J. *Handbook for Rhizobia: Methods in Legume-Rhizobium Technology* 450 (Springer, New York, 1994).
40. Ehrhardt, D. W., Atkinson, E. M. & Long, S. R. Depolarization of alfalfa root hair membrane potential by *Rhizobium meliloti* Nod factors. *Science* **256**, 998–1000 (1992).
41. Markwell, J. & Blevins, D. The minolta SPAD-502 leaf chlorophyll meter: An exciting new tool for education in the plant sciences. *Am. Biol. Teach.* **61**, 672–676 (1999).
42. Nagata, T. & Takebe, L. Cell wall regeneration and cell division in isolated tobacco mesophyll protoplasts. *Planta* **92**, 301–308 (1970).

Acknowledgements

This study was supported by IRD (JEAI_SymbiTrop); the Suranaree University of Technology (SUT), the Thailand Science Research and Innovation (TSRI), the office of national higher education science research and innovation policy council (nxpo) by Program Management Unit (PMU-B) project (B05F630107), the Agricultural Research Development Agency (ARDA) (PRP6305031340), the Franco-Thai Cooperation Program in Higher Education and Research (PHC SIAM 2017—No. 38280WF) and by the ANR grants 'SymEffectors' and 'ET-Nod' number ANR- 16-CE20-0013 and ANR-20-CE20-0012, respectively. We thank Professor Kiwamu Minamisawa for providing the *Glycine max* seeds and Dr Daniel Foncéka for providing the V. unguiculata seeds.

Author contributions

P.S., C.C., N.T., P.T. and E.G. designed the experiments. P.S., C.C., N.T., P.T., J.W., A.T., M.M., P.P., D.G., J.F., A.M.A.Z., A.C., P.T. and E.G. performed the experiments and analysed the data. E.G. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-84205-w.

Correspondence and requests for materials should be addressed to N.T. or E.G.

Reprints and permissions information is available at www.nature.com/reprints.

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021, corrected publication 2021

Article : “The Type III Effectome of the Symbiotic *Bradyrhizobium vignae* Strain ORS3257”

biomolecules

MDPI

Article

The Type III Effectome of the Symbiotic *Bradyrhizobium vignae* Strain ORS3257

Nicolas Busset [1,†], Djamel Gully [2,†], Albin Teulet [2], Joël Fardoux [2], Alicia Camuel [2], David Cornu [1], Dany Severac [3,4], Eric Giraud [2,*] and Peter Mergaert [1,*]

1 Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), F-91198 Gif-sur-Yvette, France; nicolas.busset@ird.fr (N.B.); david.cornu@i2bc.paris-saclay.fr (D.C.)
2 Laboratoire des Symbioses Tropicales et Méditerranéennes (LSTM), UMR IRD/SupAgro/INRAE/Université de Montpellier/CIRAD-Campus de Baillarguet, F-34398 Montpellier, France; djamel.gully@ird.fr (D.G.); albin.teulet@slcu.cam.ac.uk (A.T.); Joel.Fardoux@Ird.Fr (J.F.); camuel.alicia@gmail.com (A.C.)
3 Institut de Génomique Fonctionnelle, Université Montpellier, CNRS, INSERM, F-34094 Montpellier, France; dany.severac@mgx.cnrs.fr
4 Montpellier GenomiX, France Génomique, F-34094 Montpellier, France
* Correspondence: eric.giraud@ird.fr (E.G.); peter.mergaert@i2bc.paris-saclay.fr (P.M.)
† Both authors contributed equally to this article.

Abstract: Many *Bradyrhizobium* strains are able to establish a Nod factor-independent symbiosis with the leguminous plant *Aeschynomene indica* by the use of a type III secretion system (T3SS). Recently, an important advance in the understanding of the molecular factors supporting this symbiosis has been achieved by the in silico identification and functional characterization of 27 putative T3SS effectors (T3Es) of *Bradyrhizobium vignae* ORS3257. In the present study, we experimentally extend this catalog of T3Es by using a multi-omics approach. Transcriptome analysis under non-inducing and inducing conditions in the ORS3257 wild-type strain and the *ttsI* mutant revealed that the expression of 18 out of the 27 putative effectors previously identified, is under the control of TtsI, the global transcriptional regulator of T3SS and T3Es. Quantitative shotgun proteome analysis of culture supernatant in the wild type and *T3SS* mutant strains confirmed that 15 of the previously determined candidate T3Es are secreted by the T3SS. Moreover, the combined approaches identified nine additional putative T3Es and one of them was experimentally validated as a novel effector. Our study underscores the power of combined proteome and transcriptome analyses to complement in silico predictions and produce nearly complete effector catalogs. The establishment of the ORS3257 effectome will form the basis for a full appraisal of the symbiotic properties of this strain during its interaction with various host legumes via different processes.

Keywords: *Bradyrhizobium*; type III secretion system; effector; nodulation; legume symbiosis; nod factor; proteome; transcriptome

check for updates

Citation: Busset, N.; Gully, D.; Teulet, A.; Fardoux, J.; Camuel, A.; Cornu, D.; Severac, D.; Giraud, E.; Mergaert, P. The Type III Effectome of the Symbiotic *Bradyrhizobium vignae* Strain ORS3257. *Biomolecules* **2021**, *11*, 1592. https://doi.org/10.3390/biom11111592

Academic Editor: Delphine Vincent

Received: 20 September 2021
Accepted: 26 October 2021
Published: 28 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Bradyrhizobia are Gram-negative bacteria that are widespread and dominant in soils all around the world [1]. The best-known *Bradyrhizobium* strains are symbiotic nitrogen fixers, interacting with a large diversity of legumes, including plants of agronomic interest [2]. This symbiosis leads to the formation of new root organs, the nodules, in which the bacteria differentiate into a nitrogen-fixing form called bacteroids. The major part of the nitrogen fixed by the bacteroids is transferred to the plant, which can therefore grow in soils that are deficient in available nitrogen for the plant. In exchange, bacteria obtain a privileged ecological niche within the nodules where they can massively multiply thanks to nutrients provided by the plant, including reduced carbon resulting from photosynthesis.

This symbiotic process between the two partners is regulated by the exchange of signal molecules. One of the key steps allowing the initiation of symbiosis is the secretion by the bacteria of lipochitooligosaccharide signals called Nod factors (NFs), whose synthesis

is induced by the perception of plant flavonoids [3]. The recognition of NFs by the host plant leads to the activation of a signaling cascade regulating two coordinated processes, nodule organogenesis and its infection by rhizobia, thus allowing the formation of a functional nodule [4]. This Nod factor-dependent symbiotic process was considered universal for all rhizobium-legume interactions until the identification of a specific clade of photosynthetic *Bradyrhizobium* strains that lack the genes crucial for the NF synthesis but that are, nevertheless, able to establish a functional symbiosis with tropical legumes of the genus *Aeschynomene* [5]. This iconoclastic finding demonstrated the existence of non-conventional symbiotic processes.

More recently, it was shown that other *Bradyrhizobium* species can nodulate some legumes, even when their NF synthesis is abolished by mutagenesis or when the perception of NF in the host plant is inactivated. The *Bradyrhizobium elkanii* strain USDA61 can nodulate *Glycine max* cv. Enrei (soybean) as well as *Aeschynomene indica* in the absence of NFs but in a Type III Secretion System (T3SS)-dependent manner [6–8]. Similarly, the *Bradyrhizobium vignae* strain ORS3257 requires its T3SS to induce nodules on the NF-independent legume *A. indica* [7,9]. This T3SS-dependent nodulation of *A. indica* is distinct from the NF-independent process employed by the earlier described photosynthetic bradyrhizobia that produce no NFs but also have no T3SS and thus depend on a different signaling mechanism, which still remains unknown [5,10].

The T3SS was first studied in animal pathogens like bacteria of the genus *Shigella*, *Yersinia*, *Salmonella*, and *Pseudomonas*, in which it is a key pathogenesis factor [11]. The T3SS functions as a nano-syringe that enables the bacteria to inject proteins, called effectors, from its cytosol directly into neighboring host cells. The delivered effectors then interfere with cellular processes such as the cytoskeleton organization and the immune system, thereby promoting pathogenesis [12]. Similarly, phytopathogenic bacteria like *Pseudomonas*, *Xanthomonas*, *Erwinia*, and *Burkholderia* species use T3SS effectors (T3Es) to inhibit the so-called PAMP (pathogen-associated molecular pattern)-triggered immunity or PTI in plants. These T3Es are the subject of a cat-and-mouse game between pathogen and host plant as they can also be recognized by plant immune receptors (resistance (R) proteins), leading to the activation of another layer of defense, called effector-triggered immunity or ETI [13]. Interestingly, many rhizobia also possess a T3SS and use T3Es to interfere with plant immunity and increase their symbiotic efficiency and their host spectrum. However, similarly as observed for phytopathogens, some T3Es can be recognized in other hosts by R proteins triggering an ETI that blocks the symbiosis [14,15].

The NF-independent but T3SS-dependent nodulation in some soybean cultivars and in *Aeschynomene* by bradyrhizobia, suggests an additional role of the T3SS in these strains, enabling a symbiotic process in which the NF-signaling is replaced by T3SS-signaling [6–9]. A previous study, based on in silico predictions, coupling the search for *tts*-box motifs to which the TtsI regulator of rhizobial T3SS binds and the identification of homologs with known T3Es, revealed 27 putative *T3E* genes in the *B. vignae* strain ORS3257 genome [9]. They were all located within two distinct regions of the symbiotic island, one containing the nodulation (*nod*) genes and the other one containing the genes encoding the T3SS machinery (Figure 1). A broad survey of effectors in nodulating *Bradyrhizobium* strains concluded that nearly all T3Es in these bacteria are located within the symbiotic islands, suggesting a strong functional link of the T3SS with the nodulation process in these bacteria [10]. Twenty-three of the predicted T3E-encoding genes in strain ORS3257 were mutated to characterize their role during symbiosis with *A. indica* [9]. This analysis highlighted that several T3Es play roles at different stages of the symbiotic process, such as the formation of nodules, their infection, or the dampening of plant immunity to permit a chronic infection. In particular, a new effector named ErnA was identified that triggers the nodulation process.

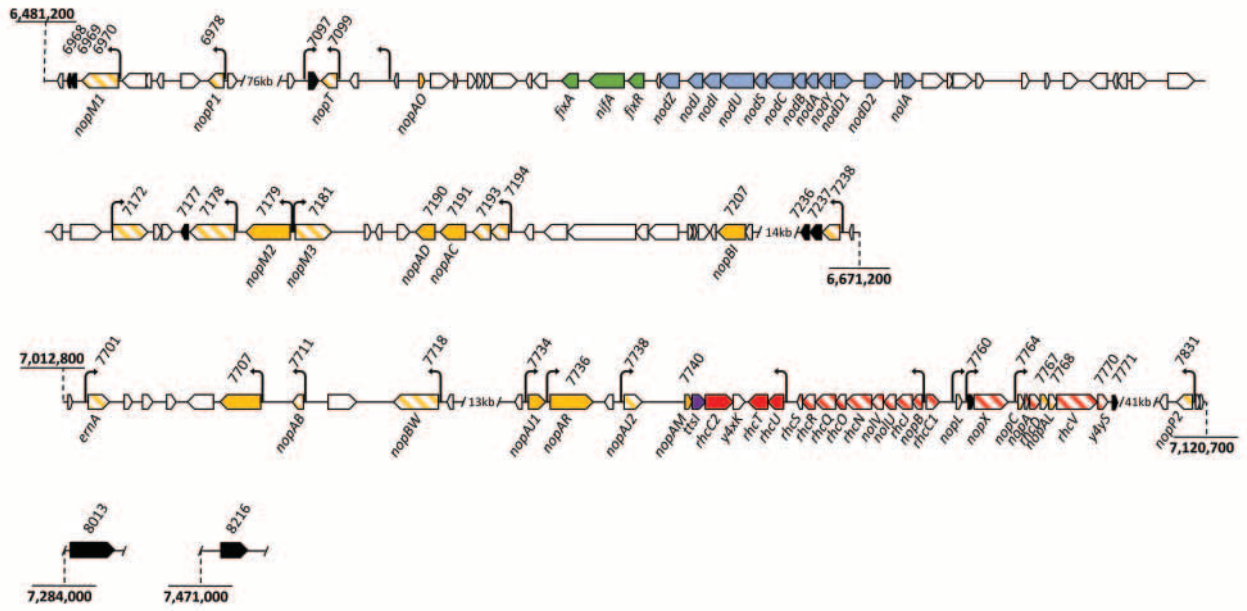


Figure 1. Genetic organization of *Bradyrhizobium* ORS3257 symbiotic island encoding the Type III secretion system and nodulation genes. Green, *nif*, and *fix* genes involved in nitrogen fixation; blue, *nod* genes; violet, *ttsI*; yellow, putative effector genes previously identified by in silico prediction; red, genes encoding components of the T3SS apparatus; black, new putative effectors; black arrows, *tts*-boxes; hatched boxes represent genes regulated by TtsI. Adapted from [9].

Since the in silico analysis revealed, on the one hand, T3E candidates with homology to known effectors but without upstream *tts*-box in their corresponding gene, and on the other hand, *T3E* genes with *tts*-box but without homology to known effectors, it is not excluded that additional T3Es are encoded in the genome of strain ORS3257 that have no *tts*-box and no homology. Furthermore, this in silico approach is only predictive and can lead to an overestimation of the number of T3Es. Therefore, revealing the complete set of effectors in strain ORS3257 with experimental methods is of cardinal importance to study the full potential of the effector arsenal, the effectome, to interfere with host signaling during nodulation. The aim of the present study was to validate and expand the T3E catalog identified in silico in strain ORS3257 by combining a transcriptome approach to identify the set of genes under the control of the TtsI regulator with a proteome analysis to determine the proteins secreted in a T3SS-dependent manner by ORS3257. This multi-omics approach experimentally validated a majority of previously predicted T3Es and identified a new one containing a ubiquitin-like protease (ULP)-like domain.

2. Materials and Methods

2.1. Bacterial Strains and Growth Conditions

The *B. vignae* strain ORS3257 and its derivatives (Table S1, see Supplementary Materials) were grown in yeast mannitol (YM) medium at 28 °C or 34 °C. *Escherichia coli* strains were grown at 37 °C in Lysogeny broth (LB) medium. When required, the media were supplemented with the appropriate antibiotics at the following concentrations: 50 µg/mL kanamycin (Km), 20 µg/mL nalidixic acid (Nal), 20 µg/mL cefotaxime (Cefo), and 20 µg/mL spectinomycin (Sp).

2.2. Plasmid Construction and Mutagenesis

All the constructs made in this study are listed in Table S1 (see Supplementary Materials), which also includes the primers and the cloning strategies. For the construction of the deletion mutants by double crossing-over, 750 to 1000 bp PCR fragments corresponding to the upstream and downstream flanking regions of the gene or locus to delete were merged by overlap extension PCR and cloned into pNPTS129, which carries the *sacB* gene. The

resulting plasmids were then transferred by mating into the ORS3257 strain, and deletion mutants were obtained by counter selection of the *sacB* gene on plates containing 10% sucrose, as described previously [16]. For the construction of insertional mutants by single crossing-over, 250 to 400-bp internal fragments of the gene to mutate were amplified by PCR and were cloned into the non-replicative plasmid pVO155-npt2-GFP-npt2-Cefo [7]. The plasmids were transferred by mating into the ORS3257 strain, resulting in their integration in the genome by single crossing-over and interruption of the targeted gene.

For the construction of the *ttsI* overexpressor strains, the complete open reading frame was cloned into the non-replicative plasmid pVO155-npt2-GFP-npt2-Cefo under the control of the *npt2* promoter [7]. The plasmid was transferred by mating into the ORS3257 wild-type (WT) and ORS3257Δ*T3SS* strains. The integration of the plasmid in the genome by a single crossing-over resulted in a merodiploid strain carrying two copies of the *ttsI* gene, one under the control of its native promoter and a second copy under the control of the constitutive and strong *npt2* promoter.

The ORS3257 WT and ORS3257Δ*T3SS* strains expressing His_6-tagged proteins were constructed by cloning a fragment encompassing the complete *BRAD3257_v2_7792*, *BRAD3257_v2_8013*, or *BRAD3257_v2_8216* gene and their upstream promoter regions using reverse primers containing a His_6-tag sequence. These inserts were then cloned into the *Xba*I or *Xho*I site of pVO155-npt2-GFP-npt2-Cefo and introduced into the ORS3257 WT and ORS3257Δ*T3SS* strains by single crossing-over. The resulting merodiploid strains carried two copies of the putative effector genes, both under the control of the native promoter, one encoding the WT protein and the other the His_6-tagged protein. All the constructed strains were verified by PCR to confirm their identity.

2.3. Transcriptome Analysis

B. vignae ORS3257 WT strain and its Ω*ttsI* mutant [9] were cultivated in BNM-B minimal medium at 34 °C [17]. When the OD_{600} reached ~0.4, 5 µM of genistein dissolved in DMSO or DMSO alone was added and the cultures were harvested after 24 h. Total RNA was extracted using the Ribopure Bacteria kit (Ambion), treated with DNase I (Qiagen), and submitted to bacterial rRNA depletion using the RiboZero kit for Gram-negative bacteria (Illumina Inc., Paris, France). Each condition was performed in independent biological duplicates. The eight corresponding RNA-seq libraries were generated and sequenced using HiSeq2000 technology (single reads 50bp; Illumina Inc., USA). For each sample, 32 to 41 million reads were obtained, out of which more than 92% were assigned to mRNA. The sequencing data of this project was submitted to the Sequence Read Archive (SRA) and obtained accession n° SRR14479858 to SRR14479861. The transcriptomic data were analyzed as previously described [18], using the ORS3257 genome as a reference [19]. The differential expression level of each gene between conditions was determined using the DESeq2 statistical analysis [20]. Quality data of the RNA-seq are shown in Figure S1 (see Supplementary Materials).

2.4. Proteome Analysis

For isolation of extracellular proteins, bacteria were grown in 600 mL of BNM-B minimal medium [17] inoculated at 1:200 with precultures of ORS3257 WT or ORS3257Δ*T3SS* strains and their derivatives overexpressing TtsI (See Table S1 in Supplemental Materials for the constructions). Bacteria were cultivated in the presence of 5 µM genistein at 28 °C and 200 rpm for 48 h. Supernatants were recovered from the bacterial cultures by two rounds of centrifugation (4000× *g* at 4 °C for 1 h and 8000× *g* at 4 °C for 30 min). Supernatants were lyophilized, and extracellular proteins were extracted as described previously [21].

For mass spectrometry analysis, proteins from each condition were successively resuspended in Rapigest (Waters), reduced/alkylated with DTT and iodoacetamide and subjected to trypsin digestion at 37 °C (overnight). After hydrolysis of Rapigest (protocol from Waters), tryptic peptide solutions were vacuum dried and resuspended in solvent A (0.1% (*v*/*v*) formic acid). Tryptic peptides from different conditions were analyzed by

nanoLC-MS/MS with the Triple-TOF 4600 mass spectrometer (ABSciex) coupled to the nanoRSLC ultra-performance liquid chromatography (UPLC) system (Thermo Scientific) equipped with a trap column (Acclaim PepMap100C18, 75 µmi.d. × 2 cm, 3 µm) and an analytical column (Acclaim PepMapRSLCC18, 75 µmi.d. × 50 cm, 2 µm, 100 Å). Peptides were loaded at 5 µL/min with 0.05% TFA in 5% acetonitrile, and peptide separation was performed at a flow rate of 300 nL/min with a 5 to 35% solvent B (0.1% formic acid in 100% acetonitrile) gradient in 40 min. NanoLC-MS/MS experiments were conducted in a data-dependent acquisition method by selecting the 20 most intense precursors for CID fragmentation with Q1 quadrupole set at low resolution for better sensitivity.

For protein identification, raw data were processed with the MS Data Converter software (AB Sciex) for generating.mgf data files, and protein identification was performed using the MASCOT search engine (Matrix science, London, UK) against the ORS3257 database with carbamidomethylation of cysteine residues set as a fixed modification and oxidation of methionine residues as a variable modification. Peptide and fragment tolerance were respectively set at 20 ppm and 0.05 Da. Proteins were validated when identified with at least two unique peptides and only peptides with Mascot ion scores above identity threshold at less than 1% FDR (false discovery rate) were considered [22]. This analysis was performed in triplicate.

Predictions of the cellular localization of proteins were carried out using the tools available on the MicroScope platform (https://mage.genoscope.cns.fr/microscope/, accessed on 27 October 2021). The ULP domain was inspected using NCBI Conserved Domain Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, accessed on 27 October 2021). Repeat domains were identied using Tandem Repeats Finder (https://tandem.bu.edu/trf/trf.html, accessed on 27 October 2021). NLSs were predicted using NLS Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi, accessed on 27 October 2021). SUMO-protease domains of BRAD3257_v2_7792, Bel2–5 (*B. elkanii* strain USDA61) and NopD (*Bradyrhizobium* sp. XS1150) were aligned using the MUSCLE algorithm.

2.5. Western Blot Analysis

Six hundred milliliters of BNM-B minimal medium were inoculated at 1:200 with precultures of ORS3257 or ORS3257Δ*T3SS* strains expressing a His_6-tagged version of the putative effectors (Table S1, see Supplementary Materials). Bacteria were cultivated in the presence of 5 µM genistein at 28 °C for 48 h at 200 rpm. Bacterial cells and exopolysaccharides were separated from the supernatant by two centrifugation steps (first step: 1 h, 4000× *g*, 4 °C; second step: 30 min, 8000× *g*, 4 °C). Proteins in the supernatant were precipitated using trichloroacetic acid as previously described [23] and resuspended in 75 µL of NuPAGE™ LDS Sample Buffer (Thermofischer, Illkirch, France) for SDS-PAGE analysis. Bacterial cells were resuspended in 5 mL of solubilization buffer (50 mM Tris-HCl pH 8, 20 mM imidazole and 300 mM NaCl) and lysed by 5 freeze-thaw cycles in liquid nitrogen. Cell debris was removed by centrifugation (30 min, 21,000× g, 4 °C), and LDS Sample Buffer 4x was added to the supernatants for SDS-PAGE analysis.

Twenty-five µL protein solutions were separated on 12.5% SDS-PAGE gels and transferred to PVDF membranes. The membranes were then blocked using PBSTM consisting of 5% non-fat milk in PBST buffer (1× PBS with 0.1% Tween 20). Anti-His_6 Mouse Antibodies (1:1000; #SAB1305538; Merck) were added to the PBSTM and the membranes were incubated 2 h at room temperature. Then, the membranes were incubated for 2 h with peroxidase-conjugated anti-mouse antibodies (1:500; #A9044; Merck). Revelation of Western blots was performed by chemiluminescence using Pierce™ ECL Plus Western Blotting Substrate (#32132; ThermoFischer) according to the manufacturer's protocols.

2.6. Symbiotic Tests

Culture and rhizobial inoculation of *A. indica* plants were performed as previously described [7]. For nodulation and the nitrogen-fixation assay, 12 plants per condition at

21 days post-inoculation (dpi) were analyzed for quantification of the number of nodules formed and for determination of the nitrogenase activity by the acetylene reduction assay, as previously described [9]. The experiments were carried out in duplicate. Microscopy analysis was performed on fresh nodules, sectioned with a Leica VT1200S vibratome (Leica Microsystems GmbH) into 30 µm tissue slices. The nodule sections were incubated 15 min in live/dead staining solution and were analyzed as previously described [7].

2.7. Outer Membrane Vesicles Observation by Transmission Electron Microscopy

Precultures grown in YM medium of the *B. vignae* strain ORS3257 and its derivatives were inoculated at 1:200 in 20 mL half-strength YM and cultures were grown at 28 °C for 3 days. Ten µL of the cultures were deposited on a microscopy grid. After 5 min, the excess liquid was removed with a filter paper and then 10 µL of 0.5% uranyl acetate was deposited for staining for one minute. After removing the liquid on the grids with filter paper, they were observed using a STEM JEOL 1400 microscope.

3. Results

3.1. Transcriptome Identification of the Genes under the Control of TtsI and the Flavonoid Genistein in Bradyrhizobium vignae ORS3257

In rhizobia, the expression of the T3SS machinery and T3Es is under the control of the TtsI transcriptional regulator, which is in turn induced by the NodD transcriptional activator and plant-derived flavonoids [24]. TtsI is activated in strain ORS3257 by the flavonoid genistein [9]. In order to identify the genes under the control of TtsI and genistein, a transcriptome analysis was carried out on in vitro cultures of ORS3257 WT and the ORS3257Ω*ttsI* mutant strain in the presence or absence of the inducer genistein.

The comparison of the gene expression between the WT and ORS3257Ω*ttsI* in the presence of genistein revealed the overexpression (fold change $\geq$ 2, p-value < 0.01) of 61 coding sequences out of 8969 in the WT strain compared to the mutant (Table S2, see Supplemental Materials). Out of these 61, 40 genes were significantly induced (fold change $\geq$ 2, p-value < 0.01) by genistein, and their genistein-induction was dependent on TtsI.

3.1.1. T3SS Machinery and Predicted T3SS Effectors Genes Regulated by TtsI

In the genome of the ORS3257 strain, 18 genes involved in the synthesis of the T3SS machinery are known and grouped within the symbiotic island [9]. Among them, 14 genes (*nopBrhcJnolUVrhcNOQRS*, *nopX*, *nopArhcDVy4yS*), located in three transcriptional units carrying a *tts*-box in the upstream region, were upregulated in the WT condition with respect to the *ttsI* mutant, confirming that their expression is under the control of TtsI and in agreement with the presence of a *tts*-box (Figures 1 and 2). In addition, we noticed that the *rhcC1* gene, which is adjacent to the *nopB* operon but transcribed in the opposite orientation, is also upregulated by genistein in a TtsI-dependent fashion (Table S2), despite the fact that the *tts*-box is oriented towards the *nopB* operon (Figure 1). This expression pattern suggests that this *tts*-box controls expression in both orientations. Twelve of these genes code for the basal body of the T3SS machinery and the three others for the needle (*nopA*, *nopB*) and the translocator (*nopX*). In contrast, the expression of the genes *rhcC2*, *rhcU*, and *rhcT*, encoding additional T3SS components, seemed not to be dependent on TtsI in the tested conditions of the transcriptome analysis (Table S3), despite the presence of a *tts*-box upstream of the *rhcUT* operon. These results confirm that the transcriptome analysis allows the identification of genes under the control of TtsI, but also suggest that the presence of a *tts*-box is not an absolute indicator of TtsI-dependent expression or that a *tts*-box can direct expression in the opposite orientation. Additional experiments using reporter gene fusions or RT-qPCR with more genistein-induction times could be used to test the latter hypotheses in more detail.

Gene ID	Gene Name	*In silico* predicted effectors [9]		Putative effector genes induced by TtsI & genistein	Putative effectors identified in secretomes of WT or OETtsI[1]
		homology	*tts* box		
Brad3257_v2_6968	/			√	
Brad3257_v2_6969	/			√	
Brad3257_v2_6970	*nopM1*	√	√	√	√
Brad3257_v2_6978	*nopP1*	√	√	√	√
Brad3257_v2_7097	/			√	
Brad3257_v2_7099	*nopT*	√	√	√	√
CDS6578292D	*nopAO*	√			
Brad3257_v2_7172	/		√	√	√
Brad3257_v2_7177	/			√	
Brad3257_v2_7178	/		√	√	√
Brad3257_v2_7179	*nopM2*	√	√		
Brad3257_v2_7181	*nopM3*	√	√	√	√
Brad3257_v2_7190	*nopAD*	√	√		
Brad3257_v2_7191	*nopAC*	√	√		
Brad3257_v2_7193	/		√	√	
Brad3257_v2_7194	/		√	√	
Brad3257_v2_7207	*nopBI*	√			
Brad3257_v2_7238	/		√	√	√
Brad3257_v2_7701	*ernA*		√	√	√
Brad3257_v2_7707	/		√		
Brad3257_v2_7711	*nopAB*	√	√	√	√
Brad3257_v2_7718	*nopBW*	√	√	√	√
Brad3257_v2_7734	*nopAJ1*	√	√		√
Brad3257_v2_7736	*nopAR*	√	√		√
Brad3257_v2_7738	*nopAJ2*	√	√	√	√
Brad3257_v2_7740	*nopAM*	√			
Brad3257_v2_7759	*nopL*	√	√	√	
Brad3257_v2_7760	/			√	
Brad3257_v2_7764	*nopC*	√	√	√	√
Brad3257_v2_7767	*nopAL*	√	√	√	
Brad3257_v2_7768	/		√	√	
Brad3257_v2_7771	/			√	
Brad3257_v2_7792	/				√
Brad3257_v2_7831	*nopP2*	√	√	√	√
Brad3257_v2_8013	/				√
Brad3257_v2_8216	/				√

[1]OETtsI: TtsI over-expressing strain

Figure 2. Overview of putative effectors in *Bradyrhizobium vignae* strain ORS3257.

Among the 27 previously identified putative effector genes, 18 are overexpressed in the WT-strain compared to the *ttsI* mutant (Figures 1 and 2), suggesting that their transcription

is under the control of TtsI. Among them, seven encode the common rhizobial effectors NopM1, NopP1, NopT, NopM3, NopL, NopC, and NopP2, four encode the *Bradyrhizobium*-specific effectors NopAB, NopBW, NopAJ2, and NopAL, and seven correspond to putative novel effectors predicted in our previous study (Figures 1 and 2). In addition, two putative effectors, encoded by *Brad3257_v2_7707* and *Brad3257_v2_7734* (*nopAJ1*), are potentially also under the control of TtsI, but their fold change or associated *p*-value fell below the selected cutoff values (Table S3). All the above putative effector genes have an upstream *tts*-box, consistent with their TtsI-dependent expression. However, the expression of four additional putative effector genes, *nopM2* (*Brad3257_v2_7179*), *nopAD* (*Brad3257_v2_7*190), *nopAC* (*Brad3257_v2_7191*), and *nopAR* (*Brad3257_v2_7736*), each possessing a *tts-box* motif in their upstream region, was not dependent on TtsI in the tested conditions (Table S3). The three remaining predicted effector genes, *nopAO* (*CDS6578292D*), *nopBI* (*Brad3257_v2_7207*), and *nopAM* (*Brad3257_v2_7740*), do not have an upstream *tts-box* motif and were found not to be regulated by TtsI (Figure 1; Tables S2 and S3, see Supplementary Materials).

3.1.2. Identification of New Putative Effectors by Transcriptome Analysis

Among the open reading frames (ORFs) overexpressed in the WT condition compared to the *ttsI* mutant, 12 are scattered all over the ORS3257 genome and are located outside of the symbiotic island (Table S2, see Supplementary Materials). Ten ORFs are conserved bacterial genes encoding transcription factors, metabolic proteins, or components of the flagellar machinery and are therefore unlikely effectors. The last two genes, *BRAD3257_v2_1638* and *BRAD3257_v2_4946* do not have homologies with known proteins, and *BRAD3257_v2_1638* is even specific for strain ORS2357. However, because of the prevalent location of effectors in the symbiotic island, these last two genes were not further considered.

All the remaining TtsI-dependent genes are located within the symbiotic island. Unexpectedly, among them are the five nodulation genes, *nodY*, *nodA*, *nodB*, *nodC*, and *nodS*, coding for the synthesis of NFs (Table S2, see Supplementary Materials). The regulation of nodulation genes in bradyrhizobia is complex, involving a network of several regulators. In *B. diazoefficiens* USDA110, for which the mechanisms of nodulation gene regulation are the best described, this includes the two NodD proteins (NodD1 and NodD2), which act antagonistically, the two-component regulatory system NodV/NodW, which activates the *nodD1* and *nodYABCS* genes and the NolA regulator which activates *nodD2* [25]. But to our knowledge, the regulation of nodulation genes by the TtsI regulator was not reported before. Our transcriptome data showed that the *nodYABCS* genes are upregulated by genistein in the WT, in agreement with the presence of a well-conserved *nod*-box upstream of *nodY* and NodD regulation [25]. In contrast, no genistein induction of these genes was observed in the *ttsI* mutant background. Thus, TtsI appears to be an additional crucial regulator of the nodulation genes in strain ORS2357, despite the absence of a *tts*-box upstream of *nodY* or the two *nodD* genes. However, the details of this regulation as well as the interplay with the other regulators of the strain require further investigation.

Seven other coding sequences (CDS) are present in the symbiotic island (Table S2, see Supplementary Materials). Among them, two genes (*Brad3257_v2_6968* and *Brad3257_v2_6969*) with a small size, 195 and 291 bp, are downstream of *nopM1*, which is preceded by a *tts*-box (Figure 1). The *Brad3257_v2_6968* gene encodes an uncharacterized protein that is conserved among many bradyrhizobia but is not found outside of this clade. The *Brad3257_v2_6969* CDS corresponds to a protein of unknown function specific to ORS3257. The CDS, *Brad3257_v2_7236* and *Brad3257_v2_7237*, correspond to fragments of the neighboring putative effector gene *Brad3257_v2_7238*, which displays homology with trehalose-6-phosphate synthase effectors of plant pathogens [26] and which was predicted by the presence of a *tts*-box and validated by our transcriptome data.

Finally, the transcriptome analysis pointed out five additional candidate T3SS effector genes (Table S2, see Supplementary Materials; Figures 1 and 2). These genes include *Brad3257_v2_7097* near *nopT* but orientated in the opposite direction and having a *tts*-

box in its promoter region, *Brad3257_v2_7177* downstream of the predicted effector gene *Brad3257_v2_7178* and possibly under the control of the same promoter, *Brad3257_v2_7760* located just upstream of the *nopX* gene and preceded by a *tts*-box, and *Brad3257_v2_7771* encoding a small protein of 52 amino acids. The latter gene is probably part of a large operon, starting with *Brad3257_v2_7764* (*nopC*) and encoding both T3SS machinery components and putative effectors.

In summary, the gene expression analysis demonstrated the TtsI-dependent expression for 18 (or 20 if *Brad3257_v2_7707* and *Brad3257_v2_7734* (*nopAJ1*) are included despite falling below the set cutoffs for significant differential expression) predicted effectors and identified up to 6 new potential effectors (*Brad3257_v2_6968*, *Brad3257_v2_6969*, *Brad3257_v2_7097*, *Brad3257_v2_7177*, *Brad3257_v2_7760*, and *Brad3257_v2_7771*) located within the symbiotic island. However, these additional putative effectors remain speculative considering that they could correspond to fragments of degenerating genes and that the largest encoded protein among them is less than 100 amino acids.

3.2. T3SS-Dependent Secretome in Bradyrhizobium vignae ORS3257

In order to identify the proteins secreted in a T3SS-dependent manner by the ORS3257 strain, a liquid chromatography coupled to mass spectrometry (LC-MS) shotgun proteome analysis was performed on the culture supernatant of bacteria grown in the presence of genistein to induce the expression of the T3SS machinery and *T3Es* genes. The analysis was carried out on the WT strain and a T3SS deletion mutant (Δ*T3SS*), as well as the corresponding strains that over-express the TtsI regulator to increase the production and secretion of effectors in the culture medium.

3.2.1. Cytosolic and Membrane Proteins Are Present in the Supernatant

The total number of proteins identified in the culture supernatants for the four conditions combined, validated with the identification of at least two unique peptides, is 563 (Figure 3a; Table S4, see Supplementary Materials). These proteins included predicted secreted proteins (101), unknown proteins without a clear, predictable localization (159), as well as many proteins predicted to be located in the membranes (152) or the cytosol (146) of the bacteria (Figure 3b). It is common to find cellular proteins in the supernatant due to incomplete removal of cells, but the centrifugation protocol used to separate the bacterial cells from the supernatant minimized this issue. An alternative source of this type of protein in the culture supernatants is the production of outer membrane vesicles (OMV) [27]. Indeed, the protocol used to remove bacterial cells from the culture supernatants does not eliminate OMVs, which are then lysed during protein precipitation of the supernatants. Observations by negative staining using a transmission electron microscope of ORS3257 and T3SS mutant cultures revealed the presence of OMVs in all cultures, indicating that at least some of the identified cellular proteins in the proteomic analyses can originate from these vesicles (Figure 3c).

Proteins commonly found in OMV of many species of bacteria such as *E. coli*, *Pseudomonas syringae*, or *Helicobacter pylori* are registered in the Vesiclepedia repository (http://microvesicles.org, accessed on 27 October 2021) [28]. The comparison of proteins present in the supernatant of the different conditions of this analysis with the proteins identified in OMV of other bacterial species showed that 182 identified membrane or cytosolic proteins could originate from OMVs. In addition, it is possible that *Bradyrhizobium*-specific proteins, which are therefore absent from this database, are also present in OMVs (Table S4, see Supplementary Materials).

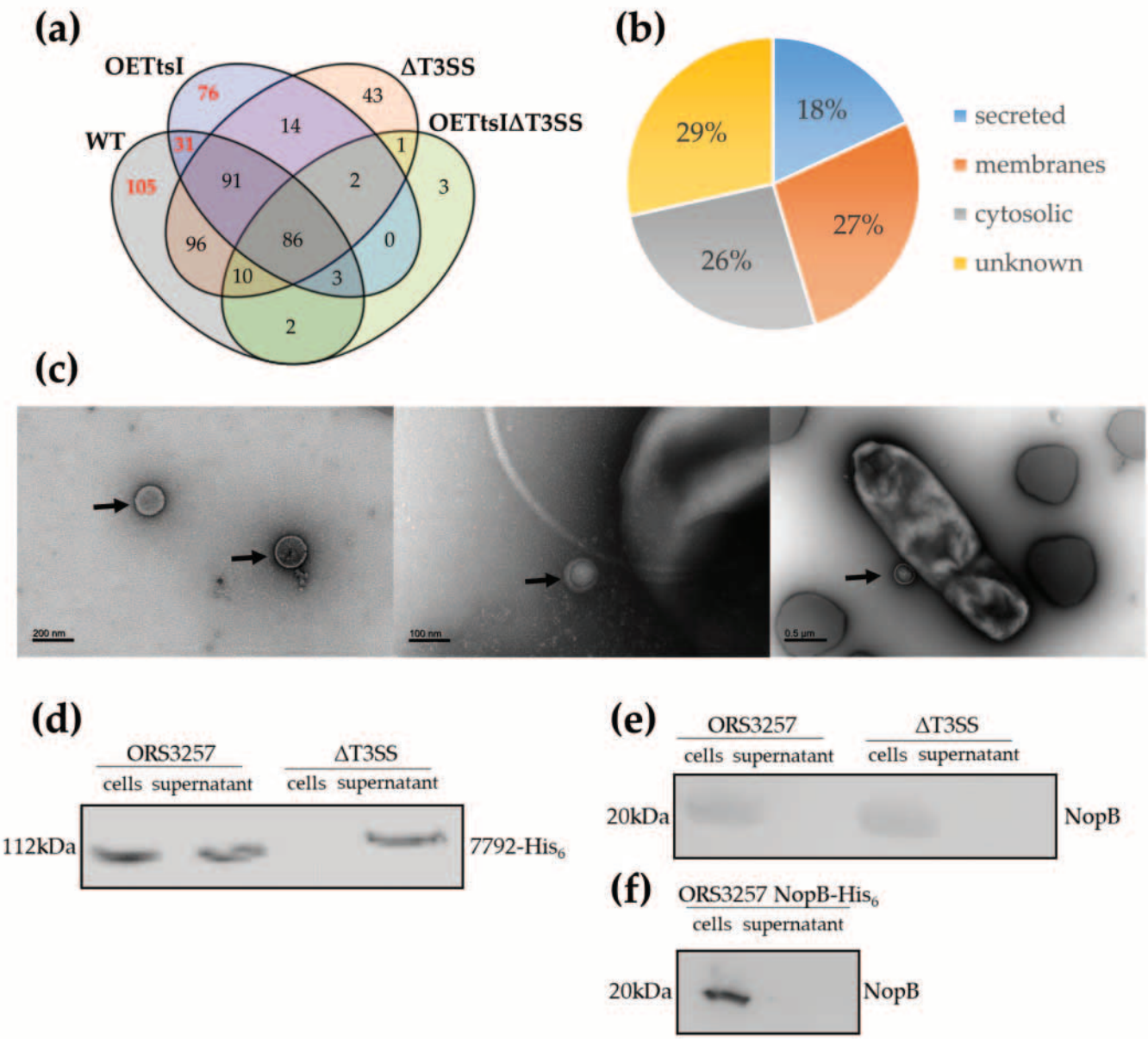


Figure 3. T3SS-dependent secretome in *Bradyrhizobium vignae* ORS3257. (**a**) Venn diagram of secreted proteins by ORS3257 WT and T3SS mutant strains. OETtsI is the ORS3257 strain overexpressing TtsI. ΔT3SS carries a deletion of the *rhcN* gene, resulting in a non-functional T3SS. In red, proteins not found in the supernatant of the T3SS mutants, which are candidate T3SS effectors. (**b**) Origin of proteins identified in the supernatant of ORS3257 and its T3SS mutant derivative. (**c**) Observation by negative staining using a transmission electron microscope of the ORS3257 WT (left and middle), and the ORS3257 OETtsI (right) strains in liquid culture. Arrows indicate OMVs. (**d**) Western blot analysis of the BRAD3257_v2_7792 effector in cell-associated and supernatant proteins of ORS3257 WT and ΔT3SS strains with an anti-His_6 antibody. (**e**) Western blot analysis of the NopB T3SS-needle protein in cell-associated and supernatant proteins of ORS3257 WT and ΔT3SS strains with an anti-NopB antibody. (**f**) Western blot analysis of NopB in cell-associated and supernatant proteins of ORS3257 WT with an anti-His_6 antibody.

3.2.2. The Expression Level of the T3SS Apparatus Modulates Protein Production and Secretion in *Bradyrhizobium vignae* ORS3257

The protein content of the supernatants of the WT and *ttsI*-overexpressing strains diverge by more than half of the proteins and only 211 are commonly present, out of a total of 516 proteins detected in the supernatants of these two strains (Figure 3a). This variability is also observed in the comparison of the supernatants of the WT strain and

the Δ*T3SS* mutant, but to a lesser extent with 283 common proteins out of 484 (Table S4, see Supplementary Materials). A large portion of the variable proteins is derived from the cytosol or bacterial membranes (Table S4, see Supplementary Materials). It is known that the T3SS synthesis and activity has an important energetic cost for the bacterial cell, hence the synthesis of the T3SS machinery and the expression of effectors are under the control of inducible promoters such as TtsI for rhizobial strains [24,29]. Therefore, we propose that the variability between conditions is an indirect effect resulting from changes in bacterial metabolism due to the energy cost of expression or overexpression of the T3SS machinery. Moreover, since TtsI is a regulatory protein, it is possible that its overexpression has unexpected effects, altering the expression of cellular processes that are not directly under its control in WT cells.

3.2.3. T3SS-Dependent Secreted Proteins by *Bradyrhizobium vignae* ORS3257 and Identification of New Putative Effectors

The candidate Type III effectors are expected to be found in the WT conditions with or without overexpression of TtsI and absent from the supernatant of the T3SS mutants. This is the case for 212 proteins, but only 31 of them are common to the WT and TtsI-overexpressing strain (Figure 3a; Table S4, see Supplementary Materials). Satisfyingly, all the candidate effectors determined by prediction and transcriptome analysis and detected in the proteome analysis were found in the supernatant of WT or the TtsI-overexpressing strain and were absent from the supernatant of the T3SS mutants (Figure 2; Table S4, see Supplementary Materials). Among the 212 proteins, 51 are membrane proteins, 58 have a cytosolic origin and 25 are secretory proteins but their secretion is predicted to be in a T3SS-independent manner. These proteins can be eliminated from candidates to be T3Es, leaving 78 proteins to be further considered (Table S4, see Supplementary Materials). Among them, besides the predicted effectors, 62 proteins whose cellular localization is unknown, were detected. However, the majority of the genes encoding these proteins are located outside the symbiotic island, excluding them from consideration.

The 18 remaining proteins, present only in the supernatant of the WT strain and its TtsI-overexpressing derivative and not in the corresponding T3SS mutants, are putative effectors (Figure 2; Table S4, see Supplementary Materials). It is useful to note that some putative effectors were only detected when TtsI is overexpressed, indicating the interest of overexpressing the transcriptional regulator to detect effectors secreted only in small quantities. Five of the 18 candidates were found not to be regulated by TtsI in the transcriptome analysis. Oppositely, 12 effectors whose gene expression is regulated by TtsI, including all the seven newly discovered TtsI-regulated candidates, were not found in the culture supernatant. Also 12 in silico predicted effectors were not detected by proteomics (Figure 2).

Among the 18 proteins detected only in WT and its TtsI-overexpressing derivative, 15 had been previously identified as putative effectors by in silico analyzes and 3 new candidates were identified: *BRAD3257_v2_7792*, *BRAD3257_v2_8013* and *BRAD3257_v2_8216* (Figure 2; Table S4, see Supplementary Materials). The transcriptome analysis indicated that these three genes have a TtsI-independent expression (Table S3, see Supplementary Materials), in agreement with the absence of *tts-boxes* in their promoter region.

3.2.4. Validation of T3SS-Effector Candidates Identified by Omics Using Western-Blotting

By combining the results of the transcriptomic and proteomic approaches as well as the data previously obtained by in silico analyzes, 37 putative effectors (regulated by TtsI, possessing a *tts*-box, secreted by the T3SS or similar to known effectors) were identified by one or more of these methods (Figure 2). The symbiotic role of effectors previously identified by bioinformatics has already been characterized [9]. In order to advance in the understanding of the mechanisms of the NF- independent, T3SS-dependent pathway, we aimed at characterizing newly identified putative effectors.

The transcriptomic and proteomic approaches have revealed ten candidates that may be new T3Es of the ORS3257 strain, but they have been validated by only one of the two

approaches (Figure 2). In order to determine whether these proteins are secreted in a T3SS-dependent manner, strains producing six of these putative effectors with a histidine tag (His_6) were constructed within the WT strain as well as the T3SS mutant. Subsequently, the presence of these proteins in the supernatant and the cellular content of ORS3257 strains was determined by Western blotting using an anti-histidine antibody.

Two proteins encoded by the genes *BRAD3257_v2_7760* and *BRAD3257_v2_8013* were not identified by Western blot both in the supernatants and in the pellets of the strains producing them, coupled with a His_6-tag (Figure S2a, see Supplementary Materials). Three other candidate proteins encoded by the genes *BRAD3257_v2_7097*, *Brad3257_v2_7177* and *BRAD3257_v2_8216* were detected in the cell pellets but not in the supernatants of both the WT and T3SS mutant strains producing His_6-tagged proteins (Figure S2a, see Supplementary Materials). This experience does not allow validating these five candidates as T3Es. In contrast, the protein encoded by the gene *BRAD3257_v2_7792* was identified in the cell pellets of the WT and T3SS mutant strains, but also specifically in the supernatant of the WT strain (Figure 3d). A control Western blot detecting the cytosolic GFP protein demonstrated the absence of cytosolic proteins in the culture supernatant due to cell lysis (Figure S2b, see Supplementary Materials). This result confirms the T3SS-dependent secretion of this protein and thus validates that *BRAD3257_v2_7792* codes for the synthesis of a T3E.

3.2.5. Secretion of T3SS-Needle Components

Besides the effectors, the needle components of the T3SS machinery are also commonly found among the secreted proteins [30–32]. Two of the predicted T3SS-needle components, common to rhizobia, were found in the ORS3257 culture supernatant. The first one, NopA, which is the major constituent of the needle (pilin), is secreted in a T3SS-dependent manner (Table S4, see Supplementary Materials). The second identified component is NopX, which is the translocon, a translocation pore positioned at the tip of the T3SS needle allowing the passage of T3Es into the host cell (Table S4, see Supplementary Materials) [33,34]. It is interesting to note that NopX is apparently secreted in a T3SS-independent manner since the protein is also present in the supernatant of the T3SS mutant (Table S4, see Supplementary Materials).

Another unexpected finding of the proteome study was the absence of detected secretion of NopB, which is one of the components of the T3SS-needle found in other phylogenetically related *Bradyrhizobium* species and *E. fredii* NGR234 (Table S4, see Supplementary Materials) [32,34,35]. Western blot analysis was used to detect NopB accumulation in the cells or supernatant of the wild strain, in order to confirm or not this surprising absence of the protein in the culture supernatant (Figure 3e). Using an anti-NopB antibody, it was possible to detect NopB in cells of the WT strain. However, the protein was not detected in the supernatant. Similarly, in a strain expressing NopB-His_6, the tagged protein was only detected in the cell fraction and was absent from the culture supernatant, thus confirming the results obtained by proteomic analysis (Figure 3f).

3.3. A Mutant in the BRAD3257_v2_7792 Effector Is Not Impaired in Symbiosis with Aeschynomene indica

The validation by two different approaches that the BRAD3257_v2_7792 protein is secreted in a T3SS-dependent manner coupled with the fact that this protein contains a SUMO-protease domain similar to other type III effectors (Figure S3, see Supplementary Materials) suggests a role for this effector during symbiosis [36]. In order to determine if this effector is important during the NF-independent symbiotic process, *A. indica* plants were inoculated with the mutant ORS3257Ω*7792*, unable to produce this protein. Observation of nodules elicited by the mutant at 21dpi on the roots of *A. indica* did not show differences with those produced by the WT strain, either in the number of nodules or their appearance (Figure 4a,b). In addition, these nodules are perfectly infected and have a similar nitrogenase activity to those produced by the WT strain (Figure 4c,d). These results indicate that the mutation of BRAD3257_v2_7792 does not affect the symbiotic capacities

of the ORS3257 strain during symbiosis with *A. indica* in the conditions of this experiment, suggesting that the BRAD3257_v2_7792 effector is dispensable during the NF-independent process.

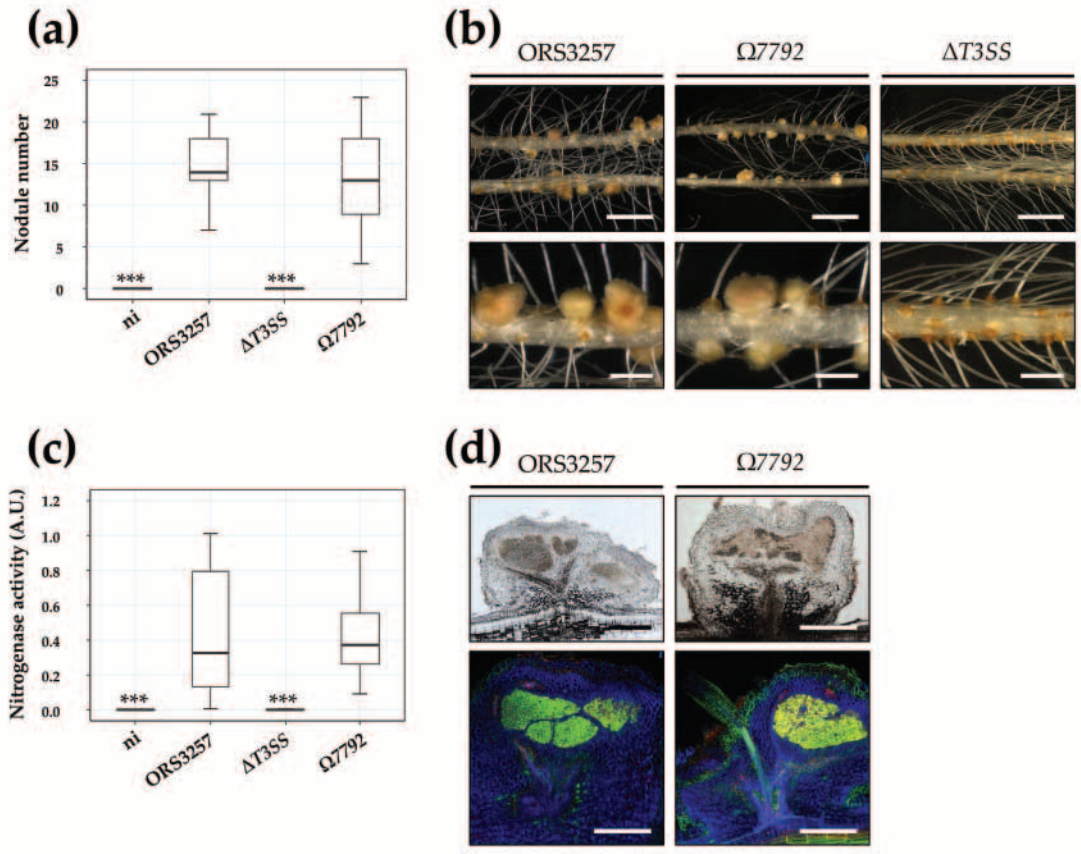


Figure 4. Symbiotic properties of the ORS3257Ω7792 mutant on *Aeschynomene indica*. (**a**) The number of nodules on *A. indica* plants at 21 days after inoculation with the WT strain ORS3257, the T3SS mutant (ΔT3SS), and the *Brad3257_v2_7792* effector mutant (Ω7792). (**b**) View of the root and the nodules induced by strain ORS3257 and the ΔT3SS and Ω7792 mutants. Scale bars, 1.5 cm (upper panels) and 2 mm (lower panels). (**c**) Nitrogen fixation activity of *A. indica* plants at 21 days after inoculation with the WT strain ORS3257 and the ΔT3SS and Ω7792 mutants. A.U., arbitrary units. Box plots in (**a**,**c**) show results of one representative experiment out of at least two independent experiments per strain (12 plants each). The central rectangle spans the first quartile to the third quartile, the bold segment inside the rectangle shows the median, and the whiskers above and below the box show the locations of the maximum and minimum value, respectively. *** $p < 0.001$, significant differences between WT ORS3257 and the non-inoculated plants (ni) or the ΔT3SS mutant using a non-parametric Kruskal–Wallis test. (**d**) Cytological analysis of the nodules induced by strain ORS3257 and the Ω7792 mutant observed by light microscopy (upper panels) or confocal microscopy after staining with Syto9 (green; live bacteria), calcofluor (blue; plant cell wall), and propidium iodide (red; plant nuclei) (lower panels). Scale bars, 500 μm.

4. Discussion

Legumes, through their symbiotic interaction with rhizobia, are of major agronomic and ecological importance. Understanding the molecular mechanisms allowing an optimal symbiosis between these two partners has been therefore of great interest for scientific research for many decades. Many rhizobial species use a T3SS to increase their host spectrum via the secretion of T3Es, sometimes called Nops for nodulation outer proteins that control immune responses in the host [37]. Characterizing the symbiotic role of T3Es could make it possible to optimize the symbiotic couples used in agronomy via a more detailed understanding of the molecular mechanisms supporting or preventing the symbiotic process. This requires, for each strain, identifying the cocktail of effectors secreted since certain effectors can act in synergy, have functional redundancies, or have

opposing effects on the symbiotic process [23,38,39]. Several studies have already been carried out in order to produce catalogs of T3Es in *Bradyrhizobium* spp. or *E. fredii* strains using in silico, transcriptomic, or proteomic approaches [10,32,40,41]. A study carried out previously revealed that the *B. vignae* ORS3257 strain, in addition to having T3Es conventionally present in other rhizobia, also has a new class of effectors involved in the direct triggering of the symbiotic process [9]. ErnA is most probably the first discovered member of this type of effector, which we propose to call ET-Nods for effectors triggering nodulation. In order to characterize the effector catalog of this strain as precisely as possible and optimize the chances of identifying new T3Es or ET-Nods, we have combined in this study transcriptomic and proteomic approaches with the previously reported in silico analysis [9].

The combined approaches resulted in the identification of 36 putative T3Es validated by at least one of the methods. It should be noted that other candidates identified by transcriptomics or proteomics were not retained in this list because being isolated in the genome outside of the symbiotic island, but they could be the object of a further study. Within this list, thirteen putative effectors (NopM1, NopP1, NopT, Brad3257_v2_7172, Brad3257_v2_7178, NopM3, Brad3257_v2_7238, ErnA, NopAB, NopBW, NopAJ2, NopC, and NopP2) were validated by all of the used methods. They have a *tts-box* motif upstream of the genes, are under the control of TtsI and are secreted in a T3SS-dependent manner. In addition, for five of them, NopM1, NopP1, NopT, NopAB, and ErnA, mutation alters the symbiotic properties of the strain during the interaction with *A. indica* [9]. These proteins, most of which have homologs in other rhizobium species, are therefore undoubted T3Es.

Other candidates are not validated by all the methods. Some candidates have only been validated by bioinformatics analysis via homology with known T3Es in other species and/or by the presence of a *tts*-box that was found to be non-functional in the transcriptome analysis. Among them are NopAO, NopM2, NopAD, NopAC, NopBI, and NopAM. Except for NopM2, for which homologs were confirmed to be secreted via the T3SS in several rhizobia (including this study for NopM1 and NopM3), to our knowledge, no homolog of the other candidate effectors was proven to be a bona fide T3E in other rhizobia. The NopAO, NopAD, NopAC, NopBI, and NopAM proteins, initially identified in *Bradyrhizobium* strains by search of the *tts*-box motif, were proposed to correspond to T3Es using a heterologous *Pseudomonas–Arabidopsi*s translocation system [42]. In this system, the N-terminus of the candidate T3E is fused to the AvrRpt2 effector of *p. aeruginosa* lacking its own N-terminal secretion signal. The hybrid protein triggers an HR in *Arabidopsis* leaves if it is translocated by *Pseudomonas* [42]. However, this approach might generate many false positives since most of the tested candidates were positive in the assay, including several components of the basal T3SS apparatus, which are not expected to be secreted. Our study indicates that NopAO, NopAD, NopAC, NopBI, and NopAM should not be considered as rhizobial T3Es. However, this conclusion does not mean that these genes do not play a possible symbiotic role considering that a *nopAO* mutant in the ORS3257 strain was found to be positively impacted during symbiosis with *A. indica* [9].

On the other hand, some candidate genes initially identified by bioinformatics, such as *nopAJ1* and *nopAR* whose expression was not confirmed to be under the control of TtsI, were found to be secreted in a T3SS-dependent manner by proteomic analysis. Our initial hypothesis that the expression of the genes involved in the synthesis of T3SS and T3Es is strictly under the control of TtsI, is therefore not accurate and we have to propose that another system for regulating the expression of effectors exists in the ORS3257 strain, as has been reported in other strains [24,43].

Certain candidates, as for example NopAL, identified by bioinformatics and expressed under the control of TtsI, were not identified by proteomics. Too low abundance or instability of these potential effectors in the supernatant or during the precipitation of proteins could explain the absence of detection of these proteins.

NopB is another undetected protein that was expected to be found in the supernatant of the T3SS-positive strains. Using three different strategies, by proteomics and by Western

blot analysis with a specific anti-NopB antibody or with a His_6-tag antibody, we could not detect the accumulation of this protein in the culture supernatant despite its detection in the cell fraction. NopB is considered as one of the major components of the T3SS needle, interacting with the NopA pilin, and classically found in the supernatant of rhizobia [32,35]. Possibly, NopB in ORS3257 is unstable in the culture supernatant, which would be in agreement with reported difficulties to detect NopB in a culture supernatant of *B. diazoefficiens* USDA110 [32,44]. Alternatively, the needle in strain ORS3257 has a different composition as compared to *E. fredii* and *B. diazoefficiens* USDA110 and, therefore, NopB has an altered and unknown function in the T3SS machinery. Besides NopB, also another anticipated needle component, NopX, behaved unexpectedly. NopX was detected in the culture supernatant, but its secretion was independent of a functional T3SS. This is unlike NopX secretion in the *E. fredii* strain NGR234, which is T3SS-dependent [29].

It is important to note that many proteins from the cytosol or membranes are present in the supernatant of the tested conditions. These proteins can hide, by their abundance, certain candidates. Our hypothesis is that these proteins can originate in part from OMVs produced by the strain since the majority of these proteins are found in the OMVs of other bacterial species [45,46]. However, it cannot be excluded that some proteins derive from cell lysis or cell leakage during the culture and the centrifugation steps. Another bias in the proteomic analysis is the modulation of proteins produced by the different mutants. In fact, a great variability is observed between the conditions in the protein content present in the supernatant, independently of the secretion by the T3SS. The regulator TtsI has been demonstrated in the NGR234 strain as regulating the synthesis of certain polysaccharides [47]. It is therefore possible that its overexpression in ORS3257 also modulates the metabolism of the bacteria and, therefore, the synthesis or export of proteins. The suppression or inactivation of the T3SS can also modulate the protein synthesis of bacteria, because the production of T3SS by bacteria has a high energy cost [48], thus explaining the variation in protein content between the tested strains.

Finally, some candidates were only identified by transcriptomics or proteomics. The production of strains producing some of these candidates with a polyhistidine tag made it possible to test by a second biochemical method the validity of these candidates. Of the six tested, only one was validated by this method. These results demonstrate that a single approach, whether biochemical or by bioinformatics, does not allow us to identify with certainty a T3E. The coupling of the different approaches made in this study thus maximizes the chances of obtaining a robust effector catalog by eliminating false positives or negatives produced by one of the used methods. Nevertheless, to conclusively qualify identified candidates as genuine T3Es, it would be necessary to use additional bio-assays to show their translocation into host cells.

BRAD3257_v2_7792, although not validated by transcriptomics, was identified by proteomics and Western blotting. In addition to its T3SS-dependent secretion, the fact that this protein contains a ubiquitin-like protease (ULP)-like domain, conserved in eukaryotic SUMO proteases and many bacterial effectors, makes it a very robust candidate for being a true T3E [49,50]. Protein desumoylation catalyzed by this type of enzyme controls cellular mechanisms ranging from transcription and cell division to ribosome biogenesis [27]. Presently, only two SUMO-protease rhizobial effectors identified in *Bradyrhizobium* strains were shown to impact positively or negatively on the symbiosis depending on the host plant [8,51,52]. One of them, the Bel2–5 effector secreted by *B. elkanii* USDA61, was also shown to be an ET-Nod since it acts as the main inducer of the T3SS-dependent symbiosis on Glycine max cv. Enrei [8,53]. Nevertheless, the *BRAD3257_v2_7792* mutant strain did not show a detectable phenotype during symbiosis with *A. indica*. However, since effectors have host-specific activities, it is likely that this effector is involved in the interaction with another host plant. Considering that *B. vignae* is also able to nodulate with a NF-dependent process other legume species such as *Vigna unguiculata* or *Arachis hypogaea* [54,55], we can speculate that this effector plays a role during the symbiotic interaction with other host legume species.

In conclusion, this study produced a catalog of candidate T3Es of the *B. vignae* ORS3257 strain and confirmed previously identified effectors whose role during symbiosis with *A. indica* and three *Vigna* species has been determined already [9,41]. The here refined catalog of effectors will allow the comprehensive analysis of the impact of the T3SS and its effectors in symbiosis with *A. indica*, or other species nodulated in a T3SS-dependent manner, or with other legume hosts such as *Vigna* spp. *or Arachis hypogea*, in order to deepen our understanding of this process. Many other strains of *Bradyrhizobium* use a T3SS-dependent symbiosis [7]. The identification and characterization of the cocktails of effectors produced by these strains will allow us to advance in the characterization of this new symbiotic pathway and perhaps to discover new ones via the identification of novel ET-Nod effectors such as ErnA and Bel2–5. Thirteen years ago, the Nod factors were considered to be essential to all rhizobium-legume symbioses, today two NF-independent symbiosis pathways have been identified. How many tomorrow?

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/biom11111592/s1, Figure S1: RNA-seq read quality; Figure S2: Western blot analysis of putative effectors in cell-associated and supernatant proteins of ORS3257 WT and ΔT3SS strains; Figure S3: Similarity between the BRAD3257_v2_7792 effector and two other rhizobial effectors containing a SUMO-protease domain; Table S1: List of strains and primers used in this study; Table S2: ORS3257 genes overexpressed in the WT condition compared to the Ω*ttsI* condition under genistein induction; Table S3: DeSeq2 analysis of the transcriptome data set; Table S4: Proteins detected in culture supernatants of ORS3257, the T3SS mutant and derivatives overexpressing TtsI.

Author Contributions: Conceptualization, N.B., D.G., A.T., E.G. and P.M.; investigation, N.B., D.G., A.T., J.F., A.C., D.C. and D.S.; writing—original draft preparation, N.B., E.G. and P.M.; writing—review and editing, D.G., A.T., E.G. and P.M. All authors have read and agreed to the published version of the manuscript.

Funding: The present work has benefited from the Imagerie-Gif core facility supported by Agence Nationale de la Recherche (ANR-11-EQPX-0029/Morphoscope; ANR-10-INBS-04/FranceBioImaging; ANR-11-IDEX-0003-02/Saclay Plant Sciences) and from the facilities and expertise of the I2BC proteomic platform (Proteomic-Gif, SICaPS) supported by IBiSA, Ile de France Region, Plan Cancer, CNRS and Paris-Sud University. N.B. benefited from a Post doc grant from the Agence Nationale de la Recherche and A.T. from a Ph.D. fellowship from the French Ministry of National Education, Higher Education and Research. The work in the E.G. and P.M. laboratories was supported by the Agence Nationale de la Recherche (ANR-16-CE20-0013 and ANR-20-CE20-0012).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The sequencing data presented in this study are openly available in the Sequence Read Archive (SRA) with reference number n° SRR14479858 to SRR14479861.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Delgado-Baquerizo, M.; Oliverio, A.M.; Brewer, T.E.; Benavent-gonzález, A.; Eldridge, D.J.; Bardgett, R.D.; Maestre, F.T.; Singh, B.K.; Fierer, N. A global atlas of the dominant bacteria found in soil. *Science* **2018**, *325*, 320–325. [CrossRef] [PubMed]
2. Ormeño-Orrillo, E.; Martínez-Romero, E. A genomotaxonomy view of the *Bradyrhizobium* genus. *Front. Microbiol.* **2019**, *10*, 1334. [CrossRef]
3. Lerouge, P.; Roche, P.; Faucher, C.; Maillet, F.; Truchet, G.; Promé, J.C.; Dénarié, J. Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* **1990**, *344*, 781–784. [CrossRef] [PubMed]
4. Oldroyd, G.E.D.; Murray, J.D.; Poole, P.S.; Downie, J.A. The rules of engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* **2011**, *45*, 119–144. [CrossRef] [PubMed]
5. Giraud, E.; Moulin, L.; Vallenet, D.; Barbe, V.; Cytryn, E.; Avarre, J.C.; Jaubert, M.; Simon, D.; Cartieaux, F.; Prin, Y.; et al. Legumes symbioses: Absence of *nod* genes in photosynthetic bradyrhizobia. *Science* **2007**, *316*, 1307–1312. [CrossRef]

6. Okazaki, S.; Kaneko, T.; Sato, S.; Saeki, K. Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17131–17136. [CrossRef]
7. Okazaki, S.; Tittabutr, P.; Teulet, A.; Thouin, J.; Fardoux, J.; Chaintreuil, C.; Gully, D.; Arrighi, J.F.; Furuta, N.; Miwa, H.; et al. Rhizobium-legume symbiosis in the absence of Nod factors: Two possible scenarios with or without the T3SS. *ISME J.* **2016**, *10*, 64–74. [CrossRef] [PubMed]
8. Ratu, S.T.N.; Teulet, A.; Miwa, H.; Masuda, S.; Nguyen, H.P.; Yasuda, M.; Sato, S.; Kaneko, T.; Hayashi, M.; Giraud, E.; et al. Rhizobia use a pathogenic-like effector to hijack leguminous nodulation signalling. *Sci. Rep.* **2021**, *11*, 1–15. [CrossRef]
9. Teulet, A.; Busset, N.; Fardoux, J.; Gully, D.; Chaintreuil, C.; Cartieaux, F.; Jauneau, A.; Comorge, V.; Okazaki, S.; Kaneko, T.; et al. The rhizobial type III effector ErnA confers the ability to form nodules in legumes. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 21758–21768. [CrossRef]
10. Teulet, A.; Gully, D.; Rouy, Z.; Camuel, A.; Koebnik, R.; Giraud, E.; Lassalle, F. Phylogenetic distribution and evolutionary dynamics of *nod* and T3SS genes in the genus *Bradyrhizobium*. *Microb. Genom.* **2020**, *6*, mgen000407. [CrossRef] [PubMed]
11. Deng, W.; Marshall, N.C.; Rowland, J.L.; McCoy, J.M.; Worrall, L.J.; Santos, A.S.; Strynadka, N.C.J.; Finlay, B.B. Assembly, structure, function and regulation of type III secretion systems. *Nat. Rev. Microbiol.* **2017**, *15*, 323–337. [CrossRef] [PubMed]
12. Mota, L.J.; Sorg, I.; Cornelis, G.R. Type III secretion: The bacteria-eukaryotic cell express. *FEMS Microbiol. Lett.* **2005**, *252*, 1–10. [CrossRef]
13. Deslandes, L.; Rivas, S. Catch me if you can: Bacterial effectors and plant targets. *Trends Plant Sci.* **2012**, *17*, 644–655. [CrossRef]
14. Tampakaki, A.P. Commonalities and differences of T3SSs in rhizobia and plant pathogenic bacteria. *Front. Plant Sci.* **2014**, *5*, 114. [CrossRef]
15. Peng, Y.; Van Wersch, R.; Zhang, Y. Convergent and divergent signaling in PAMP-triggered immunity and effector-triggered immunity. *Mol. Plant Microbe Interact.* **2018**, *31*, 403–409. [CrossRef] [PubMed]
16. Giraud, E.; Lavergne, J.; Verméglio, A. Characterization of bacteriophytochromes from photosynthetic bacteria: Histidine kinase signaling triggered by light and redox sensing. *Methods Enzymol.* **2010**, *471*, 135–159. [CrossRef]
17. Renier, A.; Maillet, F.; Fardoux, J.; Poinsot, V.; Giraud, E.; Nouwen, N. Photosynthetic *Bradyrhizobium* sp. strain ORS285 synthesizes 2-O-methylfucosylated lipochitooligosaccharides for *nod* gene–dependent interaction with *Aeschynomene* plants. *Mol. Plant Microbe Interact.* **2011**, *24*, 1440–1447. [CrossRef]
18. Gully, D.; Czernic, P.; Cruveiller, S.; Mahé, F.; Longin, C.; Vallenet, D.; François, P.; Nidelet, S.; Rialle, S.; Giraud, E.; et al. Transcriptome profiles of Nod factor-independent symbiosis in the tropical legume *Aeschynomene evenia*. *Sci. Rep.* **2018**, *8*, 10934. [CrossRef]
19. Le Quéré, A.; Gully, D.; Teulet, A.; Navarro, E.; Gargani, D.; Fardoux, J.; Cruveiller, S.; Neyra, M.; Giraud, E.; Krasova Wade, T. Complete genome sequence of *Bradyrhizobium* sp. strain ORS3257, an efficient nitrogen-fixing bacterium isolated from cowpea in Senegal. *Microbiol. Resour. Announc.* **2019**, *8*, e01449-18. [CrossRef] [PubMed]
20. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]
21. Okazaki, S.; Zehner, S.; Hempel, J.; Lang, K.; Michael, G. Genetic organization and functional analysis of the type III secretion system of *Bradyrhizobium elkanii*. *FEMS Microbiol. Lett.* **2009**, *295*, 88–95. [CrossRef] [PubMed]
22. Bonaldi, K.; Gourion, B.; Fardoux, J.; Hannibal, L.; Cartieaux, F.; Boursot, M.; Vallenet, D.; Chaintreuil, C.; Prin, Y.; Nouwen, N.; et al. 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.* **2010**, *23*, 760–770. [CrossRef] [PubMed]
23. Chang, J.H.; Urbach, J.M.; Law, T.F.; Arnold, L.W.; Hu, A.; Gombar, S.; Grant, S.R.; Ausubel, F.M.; Dangl, J.L. A high-throughput, near-saturating screen for type III effector genes from *Pseudomonas syringae*. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2549–2554. [CrossRef] [PubMed]
24. De Campos, S.B.; Deakin, W.J.; Broughton, W.J.; Passaglia, L.M.P. Roles of flavonoids and the transcriptional regulator TtsI in the activation of the type III secretion system of *Bradyrhizobium elkanii* SEMIA587. *Microbiology* **2011**, *157*, 627–635. [CrossRef] [PubMed]
25. Loh, J.; Stacey, G. Nodulation gene regulation in *Bradyrhizobium japonicum*: A unique integration of global regulatory circuits. *Appl. Environ. Microbiol.* **2003**, *69*, 10–17. [CrossRef]
26. Poueymiro, M.; Cazalé, A.C.; François, J.M.; Parrou, J.L.; Peeters, N.; Genin, S. A *Ralstonia solanacearum* type III effector directs the production of the plant signal metabolite trehalose-6-phosphate. *mBio* **2014**, *5*, e02065-14. [CrossRef] [PubMed]
27. Jan, A.T. Outer membrane vesicles (OMVs) of gram-negative bacteria: A perspective update. *Front. Microbiol.* **2017**, *8*, 1053. [CrossRef]
28. Kalra, H.; Simpson, R.J.; Ji, H.; Aikawa, E.; Altevogt, P.; Askenase, P.; Breakefield, X.; Budnik, V.; Buzas, E.; Bond, V.C.; et al. Vesiclepedia: A compendium for extracellular vesicles with Ccontinuous community annotation. *PLoS Biol.* **2012**, *10*, e1001450. [CrossRef] [PubMed]
29. Gal, J.E.; Lara-tejero, M.; Marlovits, T.C.; Wagner, S. Bacterial type III secretion systems: Specialized nanomachines for protein delivery into target cells. *Annu. Rev. Microbiol.* **2014**, *68*, 415–438. [CrossRef]

30. Marie, C.; Deakin, W.J.; Viprey, V.; Kopciñska, J.; Golinowski, W.; Krishnan, B.; Perret, X.; Broughton, W.J. Characterization of Nops, nodulation outer proteins, secreted via the type III secretion system of NGR234. *Mol. Plant Microbe Interact.* **2003**, *16*, 743–751. [CrossRef]
31. Lonjon, F.; Turner, M.; Rengel, D.; Lohou, D.; van de Kerkhove, Q.; Cazalé, A.-C.; Peeters, N.; Genin, S.; Vailleau, F. Comparative secretome analysis of *Ralstonia solanacearum* type 3 secretion-associated mutants reveals a fine control of effector delivery, essential for bacterial pathogenicity. *Mol. Cell. Proteom.* **2016**, *15*, 598–613. [CrossRef]
32. Hempel, J.; Zehner, S.; Göttfert, M.; Patschkowski, T. Analysis of the secretome of the soybean symbiont *Bradyrhizobium japonicum*. *J. Biotechnol.* **2009**, *140*, 51–58. [CrossRef] [PubMed]
33. Deakin, W.J.; Marie, C.; Saad, M.M.; Krishnan, H.B.; Broughton, W.J. NopA is associated with cell surface appendages produced by the type III secretion system of *Rhizobium* sp. strain NGR234. *Mol. Plant Microbe Interact.* **2005**, *18*, 499–507. [CrossRef] [PubMed]
34. Saad, M.M.; Staehelin, C.; Broughton, W.J.; Deakin, W.J. Protein-protein interactions within type III secretion system-dependent pili of *Rhizobium* sp. strain NGR234. *J. Bacteriol.* **2008**, *190*, 750–754. [CrossRef]
35. Saad, M.M.; Kobayashi, H.; Marie, C.; Brown, I.R.; Mansfield, J.W.; Broughton, W.J.; Deakin, W.J. NopB, a type III secreted protein of *Rhizobium* sp. strain NGR234, is associated with pilus-like surface appendages. *J. Bacteriol.* **2005**, *187*, 1173–1181. [CrossRef] [PubMed]
36. Hotson, A.; Chosed, R.; Shu, H.; Orth, K.; Mudgett, M.B. *Xanthomonas* type III effector XopD targets SUMO-conjugated proteins *in planta*. *Mol. Microbiol.* **2003**, *50*, 377–389. [CrossRef] [PubMed]
37. Staehelin, C.; Krishnan, H.B. Nodulation outer proteins: Double-edged swords of symbiotic rhizobia. *Biochem. J.* **2015**, *470*, 263–274. [CrossRef] [PubMed]
38. Kvitko, B.H.; Park, D.H.; Velásquez, A.C.; Wei, C.F.; Russell, A.B.; Martin, G.B.; Schneider, D.J.; Collmer, A. Deletions in the repertoire of *pseudomonas syringae* pv. tomato DC3000 type III secretion effector genes reveal functional overlap among effectors. *PLoS Pathog.* **2009**, *5*, e1000388. [CrossRef]
39. van Engelenburg, S.B.; Palmer, A.E. Imaging type-III secretion reveals dynamics and spatial segregation of *Salmonella* effectors. *Nat. Methods.* **2010**, *7*, 325–330. [CrossRef]
40. Rehman, H.M.; Cheung, W.L.; Wong, K.S.; Xie, M.; Luk, C.Y.; Wong, F.L.; Li, M.W.; Tsai, S.N.; To, W.T.; Chan, L.Y.; et al. High-throughput mass spectrometric analysis of the whole proteome and secretome from *Sinorhizobium fredii* strains CCBAU25509 and CCBAU45436. *Front. Microbiol.* **2019**, *10*, 2569. [CrossRef]
41. Songwattana, P.; Noisangiam, R.; Teamtisong, K.; Prakamhang, J.; Teulet, A.; Tittabutr, P.; Piromyou, P.; Boonkerd, N.; Giraud, E.; Teaumroong, N. Type 3 secretion system (T3SS) of *Bradyrhizobium* sp. DOA9 and its roles in legume symbiosis and rice endophytic association. *Front. Microbiol.* **2017**, *8*, 1810. [CrossRef] [PubMed]
42. Kimbrel, J.A.; Thomas, W.J.; Jiang, Y.; Creason, A.L.; Thireault, C.A.; Sachs, J.L.; Chang, J.H. Mutualistic co-evolution of type III effector genes in *Sinorhizobium fredii* and *Bradyrhizobium japonicum*. *PLoS Pathog.* **2013**, *9*, e1003204. [CrossRef]
43. Nguyen, H.P.; Ratu, S.T.N.; Yasuda, M.; Göttfert, M.; Okazaki, S. InnB, a novel type III effector of *Bradyrhizobium elkanii* USDA61, controls symbiosis with vigna species. *Front. Microbiol.* **2018**, *9*, 3155. [CrossRef]
44. Süß, C.; Hempel, J.; Zehner, S.; Krause, A.; Patschkowski, T.; Göttfert, M. Identification of genistein-inducible and type III-secreted proteins of *Bradyrhizobium japonicum*. *J. Biotechnol.* **2006**, *126*, 69–77. [CrossRef]
45. Altindis, E.; Fu, Y.; Mekalanos, J.J. Proteomic analysis of *Vibrio cholerae* outer membrane vesicles. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E1548–E1556. [CrossRef]
46. Sidhu, V.K.; Vorhölter, F.J.; Niehaus, K.; Watt, S.A. Analysis of outer membrane vesicle associated proteins isolated from the plant pathogenic bacterium *Xanthomonas campestris* pv. campestris. *BMC Microbiol.* **2008**, *8*, 87. [CrossRef]
47. Marie, C.; Deakin, W.J.; Ojanen-Reuhs, T.; Diallo, E.; Reuhs, B.; Broughton, W.J.; Perret, X. TtsI, a key regulator of *Rhizobium* species NGR234 is required for type III-dependent protein secretion and synthesis of rhamnose-rich polysaccharides. *Mol. Plant Microbe Interact.* **2004**, *17*, 958–966. [CrossRef] [PubMed]
48. Song, M.; Sukovich, D.J.; Ciccarelli, L.; Mayr, J.; Fernandez-Rodriguez, J.; Mirsky, E.A.; Tucker, A.C.; Gordon, D.B.; Marlovits, T.C.; Voigt, C.A. Control of type III protein secretion using a minimal genetic system. *Nat. Commun.* **2017**, *8*, 14737. [CrossRef] [PubMed]
49. Roden, J.; Eardley, L.; Hotson, A.; Cao, Y.; Mudgett, M.B. Characterization of the *Xanthomonas* AvrXv4 effector, a SUMO protease translocated into plant cells. *Mol. Plant Microbe Interact.* **2004**, *17*, 633–643. [CrossRef]
50. Tsurumaru, H.; Hashimoto, S.; Okizaki, K.; Kanesaki, Y.; Yoshikawa, H.; Yamakawa, T. A putative type III secretion system effector encoded by the MA20_12780 gene in *Bradyrhizobium japonicum* Is-34 causes incompatibility with Rj4 genotype soybeans. *Appl. Environ. Microbiol.* **2015**, *81*, 5812–5819. [CrossRef]
51. Xiang, Q.W.; Bai, J.; Cai, J.; Huang, Q.Y.; Wang, Y.; Liang, Y.; Zhong, Z.; Wagner, C.; Xie, Z.P.; Staehelin, C. NopD of *Bradyrhizobium* sp. XS1150 possesses SUMO protease activity. *Front. Microbiol.* **2020**, *11*, 386. [CrossRef] [PubMed]
52. Faruque, O.M.; Hiwa, H.; Uasuda, M.; Fuji, Y.; Kneko, T.; Sato, S.; Okazaki, S. Identification of *Bradyrhizobium elkanii* genes involved in incompatibility with soybean plants carrying the *Rj4* allele. *Appl. Environ. Microbiol.* **2015**, *81*, 6710–6717. [CrossRef] [PubMed]
53. Ratu, S.T.N.; Hirata, A.; Kalaw, C.O.; Yasuda, M.; Tabuchi, M.; Okazaki, S. Multiple domains in the rhizobial type III effector Bel2-5 determine symbiotic efficiency with soybean. *Front. Plant Sci.* **2021**, *12*, 689064. [CrossRef] [PubMed]

54. Krasova-Wade, T.; Ndoye, I.; Braconnier, S.; Sarr, B.; De Lajudie, P.; Neyra, M. Diversity of indigeneous bradyrhizobia associated with three cowpea cultivars (*Vigna unguiculata* (L.) Walp.) grown under limited and favorable water conditions in Senegal (West Africa). *Afr. J. Biotechnol.* **2003**, *2*, 13–22. [CrossRef]
55. Grönemeyer, J.L.; Hurek, T.; Bünger, W.; Reinhold-Hurek, B. *Bradyrhizobium vignae* sp. nov., a nitrogen-fixing symbiont isolated from effective nodules of *Vigna* and *Arachis*. *Int. J. Syst. Evol. Microbiol.* **2016**, *66*, 62–69. [CrossRef] [PubMed]

Annexe 2 : Matériels et méthodes complémentaires du chapitre 3

1.1. Matériels biologiques et conditions de culture

L'ensemble des souches bactériennes utilisées dans cette étude (*Bradyrhizobium*, *Agrobacterium rhizogenes* et *Escherichia coli*) ont été cultivées dans les mêmes conditions que décrites dans l'article 1. De même, la culture des plantes d'*A. indica*, *A. evenia* et les tests symbiotiques ont été réalisés comme décrits dans l'article 2.

1.2. Constructions des plasmides et complémentation

1.2.1. Ajout des tags et complémentation du mutant *ORS3257ΔernA*

L'ajout des tags 6His et Flag en C-terminal d'ErnA a été effectué par amplification de la région promotrice d'*ernA* ainsi que de sa séquence codante avec l'amorce 'GCAACCTCTAGAACTAGTCGACCGGCACCAGAACGTCTTTG' et l'amorce 'CATGTTCTAGACTAGTGGTGGTGGTGGTGGTGTGGATGATGGATGAGCCGAACGTCATTG' contenant le tag 6His ou avec l'amorce 'GCCCCTCTAGACTACTTGTCGTCATCGTCTTTGTAGTCTGGATGATGGATGAGCCGAACGTCATTG' contenant le tag Flag. Les fragments PCR ont ensuite été clonés dans le vecteur pVO155-npt2-gfp au niveau du site de restriction XbaI. Pour la fusion en C-terminal du Turbo-ID, le fragment contenant ErnA-TurboID a été amplifié en utilisant la Taq PCR Phusion à partir du pENTR221-*ernA-Turbo-ID* (envoyé par nos collaborateurs) en utilisant les amorces 'ATGGACCCATTTAACCGCATCAACCCATTC' et 'AGAAATCTAGATCATTTATCGTCATCGTCTTTGTAGTCCTTTTC'. Par ailleurs, la région promotrice d'*ernA* a été amplifiée à partir de l'ADN génomique d'ORS3257 avec les amorces 'GCAACCTCTAGAACTAGTCGACCGGCACCAGAACGTCTTTG' et 'GAATGGGTTGATGCGGTTAAATGGGTCCAT'. Les deux fragments ainsi obtenus ont été fusionnés par PCR overlap puis introduits dans le pVO-npt2-kn-gfp. L'ensemble de ces constructions a ensuite été introduit dans le mutant ORS3257*ΔernA* par simple crossing-over pour obtenir les formes complémentées associées. Les autres variants d'ErnA tagués en N-terminal dans le vecteur pJCV51 ont été obtenues par nos collaborateurs à Toulouse. Le plasmide pJCV51 utilisé est le même que précédemment utilisé dans l'article 1.

1.2.2. Création du système inductible pDEX-ernA

La région du plasmide pJCV51-*ernA* (Teulet et al., 2019) contenant la séquence codante d'*ernA* et le gène marqueur *mRFP* avec son promoteur, a été amplifiée en utilisant la Taq PCR Phusion avec les amorces 'ATGGACCCATTTAACCGCATCAAC' et 'CAAGCTCAAGCTAAGCTTCTGCAG', puis purifiée avec le kit QIAquick PCR Purification de Qiagen. Le produit de purification a ensuite été cloné dans le vecteur

d'entré pCR8-GW-TOPO (Thermo Fisher). La construction a ensuite été insérée dans *E. coli* TOP10 par électroporation. Après purification du plasmide, l'étape de recombinaison par la LRclonase II (Invitrogen) a été réalisée dans le vecteur de destination pDEX-GW-SBP (fourni par Laurent Deslandes) après linéarisation du vecteur d'entrée par XbaI, car portant la même résistance antibiotique (spectinomycine). Après transformation du produit de recombinaison dans *E. coli* DH10β, les colonies positives ont été testées par PCR puis par séquençage du plasmide. Le plasmide *pDEX-ernA-mRFP-SBP* ainsi obtenu a été introduit par électroporation dans la souche Arqua1 d'*A. rhizogenes*.

1.2.3. Obtention du mutant dans le site catalytique du domaine SUMO-protéase de Sup3

En utilisant le plasmide pVO155-*pm-sup3::sup3*-WSM1744 obtenu dans l'article 1, la forme mutée de Sup3 dans laquelle le résidu aspartate (D_{1456}) du site catalytique a été substitué par un résidu alanine ($D_{1456}A$) a été obtenue en utilisant le kit de mutagénèse dirigée Q5 (NEB) avec les amorces 'GCCTATCACTACGACTCTGTCCGGGGACAC' et 'GCCTATCACTACGCCTCTGTCCGGGGACAC'. Cette construction a ensuite été utilisée pour complémenter le mutant WSM1744*Δsup3* déjà disponible (Camuel et al., 2023).

1.3. Analyse RNA-seq

Après induction du système pDEX à la dexaméthasone (1 µM) sur des plantes transformées d'*A. evenia* âgées de 18 jours, le système racinaire de 9 plantes par condition à différents temps (0, 10, 24, 48 et 96 heures après induction) a été groupé puis immédiatement plongé dans de l'azote liquide. Ces prélèvements ont été répétés trois fois pour chacune des conditions et des temps. Trois expériences indépendantes ont été réalisées pour la condition *pDEX-ernA* contre une seule pour la condition *pDEX-eGFP* réalisée uniquement aux temps 0 et 24 heures après induction. Les échantillons ont ensuite été broyés un à un jusqu'à obtenir une poudre dont l'ensemble a été utilisée pour l'extraction et la purification des ARN selon le protocole renseigné par le fournisseur (Qiagen : RNesay Plant Mini Kit, RNase-Free DNase Set et RNeasy MinElute Cleanup Kit). La qualité des ARN obtenus et leur concentration ont été évaluées au NanoDrop 2000/200c (Thermo Fisher) et au Bioanalyseur Agilent 2100 selon les recommandations du fournisseur.

La banque de cDNA a été construite et séquencée comme décrit dans Quilbé et al., 2021 (Toulouse, France). Le reste de l'analyse a ensuite été réalisée comme décrit dans l'article 2 à l'exception du rapport log2 de FC ≥ 2.

1.4. Autres matériels et méthodes

L'ensemble des constructions et méthodes utilisées pour les expérimentations de marquage de proximité par Turbo-ID ainsi que pour les divers tests d'interaction protéiques (double-hybride en levure et Co-IP) et les tests d'activité protéase de Sup3, ont été réalisées et conduites par nos collaborateurs faisant partie du projet ANR "ET-Nod" et ne seront donc pas décrites ici. La construction contrôle *pDEX-NLS-eGFP* utilisé dans l'analyse RNA-seq a été réalisée à Toulouse avant d'être électroporée dans *A. rhizogenes* Arqua1.

.