Bases génétiques et forces sélectives impliquées dans l'évolution de nouveaux symbiotes de légumineuses

Ginaini Grazielli Doin de Moura

THÈSE

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

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Ginaini Grazielli DOIN DE MOURA

Le 9 janvier 2023

Bases génétiques et forces sélectives impliquées dans lévolution de nouveaux symbiotes de légumineuses

---

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

Spécialité : **Developpement des plantes, interactions biotiques et abiotiques**

Unité de recherche :
LIPME - Laboratoire des Interactions Plantes-Microbes-Environnement

Thèse dirigée par
Delphine CAPELA et Philippe REMIGI

Jury

Mme Claire PRIGENT-COMBARET, Rapporteure
M. Benoit ALUNNI, Rapporteur
M. Sébastien CUNNAC, Examinateur
M. Matthieu ARLAT, Examinateur
Mme Delphine CAPELA, Directrice de thèse
M. Philippe REMIGI, Co-directeur de thèse

Table des matières

Chapter 1 – Introduction

Chapter 2 - A selective bottleneck during host entry drives the evolution of new legume symbionts

Chapter 3 - Functional convergences in adaptive evolution transforming a plant pathogen into legume symbionts

Chapter 4 - Continuation of the evolution experiment did not lead to mutualism but evidenced a strong mutational convergence in the RSc2277 gene

Abbreviations

AHL:	N-Acyl Homoserine Lactone
ARA:	Acetylene Reduction Assay
CI:	Competitive Index
CSSP:	Common Symbiotic Signaling Pathway
EE:	Experimental Evolution
EPR3:	Exopolysaccharide Receptor 3
EPS:	Exopolysaccharide
ETI:	Effector Triggered Immunity
HGT:	Horizontal Gene Transfer
HT:	Horizontal Transfer
ICE:	Integrative and Conjugative Elements
IRLC:	Inverted Repeat-Lacking Clade
IS:	Insertion Sequence
IT:	Infection Thread
KPS:	Capsular Polysaccharide
LPS:	Lipopolysaccharide
LysM RLK:	Lysin Motif Receptor-Like Kinase
MAMP:	Microbe-Associated Molecular Pattern
MTI:	MAMP-Triggered Immunity
NCR:	Nodule Cystein-Rich
NF:	Nod Factor
NFP:	Nod Factor Perception
OD:	Optical Density
3OH-MAME:	Methyl 3-Hydroxymyristate
3OH-PAME:	3-Hydroxy Palmitic Acid Methyl Ester
PR:	Pathogenesis-Related (PR) Defense Proteins
PRR:	Pattern Recognition Receptor
QS:	Quorum Sensing
ROS:	Reactive Oxygen Species
SIM:	Stress Induced Mutagenesis
SNP:	Single Nucleotide Polymorphism
T2SS:	Type II Secretion System
T3E:	Type III Effector
T3SS:	Type III Secretion System
T4SS:	Type IV Secretion System
T6SS:	Type VI Secretion System
TCA:	Tricarboxylic Acid Cycle

List of Tables

Table of Figures

Chapter I

Chapter II

Acknowledgments

In 2023, 13 years are completed since I was accepted to join the microbial ecology group at the Federal University of Santa Catarina (UFSC, 2010). Since then, I had the honor of living very enriching experiences in the laboratories of soil microbiology (UFSC, 2011- 2014 and 2015-2016), of tree legumes (Embrapa agrobiology, 2012), in the group of genomics and evolution of rhizobia (Laboratoire des Interactions Plante-Microorganismes - LIPM, 2015), soil microbiology and biochemistry (Federal University of Lavras - UFLA, 2016) plant growth-promoting bacteria (UFLA, 2017 - 2019) and finally, in the SymEvol group (Laboratoire des Interactions Plantes-Microbes-Environment - LIPME, 2019-2023). Such a long path full of beautiful experiences cannot be walked alone. Therefore, science also allowed me to meet wonderful people who taught me, corrected me, opened doors for me, said "no" at the right time, people who supported me in good and bad moments and who made the whole path much more worthwhile.

First of all, I would like to express my great gratitude to my thesis director, Delphine Capela for everything I was taught, for her patience, for the discussions, for her valuable guidance. It is beautiful and inspiring to see her working with such dedication and love for what she does. On a personal level, I thank you and Laurent Sauviac for opening the doors of your home when delays with my visa prevented me from renting an apartment. I cherish all the advice you gave me as well as the moments when I felt welcome in your home.

Similarly, I thank my co-director, Philippe Remigi, for all his patience and all the help with the statistical analyses, graphs and population analysis during the thesis. Personally, I thank you for your guidance, your attention, your exemplary kindness, and your positivity. It is admirable the way you have good scientific ideas, your way of thinking and your passion for discussing science.

I thank Dr. Catherine Masson-Boivin for accepting me as a trainee in 2015 in the same laboratory where I developed my thesis. I must also thank her for all the advice at the beginning of my doctorate and for all the help during the conceptualization of this thesis.

I thank a person that I consider a friend, Saida Mouffok, for all the scientific help in the construction of the mutants used in this thesis, but also for so many good discussions, for the emotional support, for the moments when we laughed together (and there were so many

laughs), for the moments when she took care of my nutrition and brought me salads so that I could continue the experiments without delays, for so many gestures that may seem small, but were immense and very important to me. I will never forget your kindness, the way you welcomed me and always wanted to help me, and your smiling "bonjour Ginaini, ça va?" Sincerely, thank you very much.

I thank the members of the SymEvol team that I had the honor of working with, Lukas and Finlay, for helping me right at the beginning of my PhD, for the good times and for the moments that they participated in the experiments. I apologize for the huge harvests of nodules, Finlay, and for the huge number of tubes and plants, Lukas. Thank you very much!

I thank the interns that I mentored during the thesis, Marvin and Clarisse. I learned a lot from you. Thank you very much for all the help during the experiments!

I thank the newest members of the SymEvol team, Anne-Claire and Bryan, for the good times, the "coffee breaks", all the visiting and travel tips Anne-Claire gave me, and the discussions with Bryan. It was a pleasure to share the end of this experience with you.

I thank the colleagues at LIPME with whom I was able to share these three years and who were always available when I needed help. My special thanks to the teams responsible for the services available at the LIPME (laundromat and media preparation team and cleaning team), as well as to the imagerie and Genotoul platforms, which give us all the support we need to develop our work. Thank you very much for the valuable work you do.

I thank the researchers who agreed to evaluate my thesis, Dr. Claire Prigent-Combaret, Dr. Benoît Alunni, Pr.Matthieu Arlat and Dr. Sébastien Cunnac, for the time and attention they devoted to reading this work.

I thank the INRAE and the LIPME for providing all the necessary structure for the development of my thesis.

I thank the « Ministère de l'Enseignement Supérieur et de la Recherche » (MESR) for funding my Ph.D.

Personally, I thank my parents, Gilberto de Moura and Ana Paula Goulart Doin, for all the help and for allowing me to make my choices in total freedom, for being an example of honesty and kindness. Thank you very much for the many good moments we lived, for

transforming our home into a pleasant environment of unity and partnership. Thank you so much for all you have taught me by being a good example. I could not ask for better parents.

I thank my sister, Giovana Cristina Doin de Moura, my life partner, with whom I can share everything. Thank you very much for our companionship, for so many times you helped me, for so many good moments we lived, for giving me strength, joy and for our complicity during all this time.

To the great friends that science gave me, Lundoi Lee, Aline Vieira de Barros, Aline Oliveira Silva and Adriano Dorigan for the beautiful path that we have walked together, for our friendship, for calling me almost every day during the quarantine to know how I was doing and so that I would not feel lonely.

I thank Xavier Gibert for all that we have shared. I could not find a word that was enough to describe how much you have made my “toulousain” days happier and how much I appreciate our partnership and complicity. Thank you so much for going through the thesis with me, for the care and for how welcome you make me feel in foreign lands. Thank you for choosing to become my best friend.

I thank my former supervisors, Paulo Emilio Lovato and Claudio Roberto Soares, who taught me so much and with whom I could share unforgettable moments, whether in the laboratory or in the field selecting rhizobia throughout Brazil. You were examples of scientists, but also of human beings and you played a very important role during this trajectory. I repeat what I have already said to you one day: thank you because, even though I moved away, I kept knocking on your door and you always opened it with a smile on your face.

I thank the Brazilian friends and friends of other nationalities that I made in France, with whom I shared good moments and great trips. Thank you for so many smiles and so many good memories that I keep with me!

To all those who participated, as friends or colleagues, in this journey that brought me to this moment,

Thank you very much!

Abstract

Rhizobia form a polyphyletic group of symbiotic bacteria able to interact with legumes, fixing nitrogen to the benefit of the host plant. These bacteria emerged from repeated and independent events of acquisition of essential symbiotic genes, conferring their ability to nodulate and fix nitrogen. However, these transfers are not always sufficient to convert a recipient bacterium into an efficient legume symbiont. In these cases, additional adaptation steps occurring under plant-mediated selection pressure are necessary to activate and optimize the acquired symbiotic potential. In order to study the host-mediated selective forces driving the evolution of legume symbionts and to identify bacterial biological functions necessary or deleterious to symbiosis, the emergence of a new rhizobium genus has been replayed under laboratory conditions, using the plant pathogen *Ralstonia solanacearum* and the natural symbiont of *Mimosa pudica*, *Cupriavidus taiwanensis*, as receptor and donor of the symbiotic genes, respectively. After massive inoculation of *M. pudica*, three *Ralstonia* nodulating variants were trapped by the plants and were then subjected to serial cycles of nodulation. These cycles rapidly selected bacterial variants with improved symbiotic capacity. We analyzed the dynamics of molecular and phenotypic changes of populations evolved for 35 cycles. The fast adaptation was underpinned by the sequential fixation of mutational cohorts (*i.e*. mutations having similar frequency trajectories over time) within evolved populations. Identification and phenotypic characterization of beneficial mutations in these cohorts showed that adaptation was driven by the dominance of selection for host entry competitiveness over within-host proliferation. Based on gene annotations, beneficial mutations affect poorly or uncharacterized genes. Functional characterization of these mutants evidenced that the most adaptive mutations for symbiosis led to a decrease in the pathogenicity of the *Ralstonia* ancestral strain and are pleiotropic, affecting EPS production, motility, metabolism and stress resistance. Mutualism was not obtained in this experiment. However, one gene, convergently mutated in several independent evolved lineages, was found associated with a better proliferation of bacteria within nodules and the detection of low levels of nitrogenase activity.

Résumé

Les rhizobia sont des bactéries phylogénétiquement diverses, capables d'interagir avec les légumineuses et de fixer de l'azote au profit de la plante hôte. Ces bactéries ont émergé *via* des transferts horizontaux répétés et indépendants de quelques gènes essentiels pour la symbiose, conférant leur capacité d'inciter la formation d'un nouvel organe chez la plante hôte, le nodule, et de fixer de l'azote. Cependant, ces transferts ne sont pas toujours suffisants pour rendre symbiotique une bactérie receveuse. Dans ces cas, des étapes additionnelles d'adaptation sous pression de sélection de la plante sont nécessaires pour l'activation ou optimisation du potentiel symbiotique acquis. Une expérience d'évolution visant à comprendre l'émergence de nouveaux symbiotes a été initiée au laboratoire pour essayer d'obtenir un nouveau genre de rhizobium. Le phytopathogène *Ralstonia solancearum* et le symbiote naturel de *Mimosa pudica*, *Cupriavidus taiwanensis*, ont été utilisés comme receveur and donneur des gènes symbiotiques, respectivement. Après l'inoculation massive des plantes de *M. pudica*, trois variants nodulants ont été obtenus et soumis à des cycles de nodulation sériés sur *M. pudica*. Ces cycles ont sélectionné rapidement des variants ayant des capacités symbiotiques améliorées. La dynamique d'adaptation à la symbiose des populations évoluées pendant 35 cycles et les changements moléculaires ont été analysés. L'adaptation rapide des bactéries à la symbiose a été supportée par la fixation séquentielle des cohortes mutationnelles (*i.e.* mutations ayant des trajectoires temporelles de fréquences similaires au cours de l'expérience) dans les populations évoluées. L'identification et la caractérisation des mutations adaptatives dans ces cohortes suggèrent que l'adaptation des bactéries dans ce système est dirigée de façon prédominante par la sélection des bactéries pour leur compétitivité à l'entrée dans l'hôte en comparaison de leur capacité à proliférer dans l'hôte. L'annotations des gènes affectés par les mutations adaptatives identifiées indiquent qu'ils sont peu ou pas-caractérisés. La caractérisation fonctionnelle de ces mutants a démontré que plusieurs mutations adaptatives pour la symbiose diminuent la pathogénicité de la souche ancestrale de *Ralstonia* et sont pléiotropiques, agissant sur la production d'EPS, la motilité, le métabolisme et la résistance au stress. Le mutualisme n'a pas été obtenu dans cette expérience d'évolution. Cependant, un gène, muté de façon convergente dans différentes lignées évoluées, a été associé à une prolifération améliorée des bactéries dans les nodules et la détection de faibles niveaux d'activité de l'enzyme nitrogénase.

CHAPTER 1

Introduction

1. Rhizobia: the N_2-fixing symbionts of legumes

1.1. Nitrogen, an essential component for plant growth

Nitrogen is an essential component in living cells, especially for the biosynthesis of amino acids, nucleic acids, coenzymes, vitamins, pigments and chlorophyll (Malavolta, 1980; Taiz and Zeiger, 2014). Even if this nutrient corresponds to 78% of the atmosphere, it is not always present in a chemical form available to all organisms. Around 99.96% of the nitrogen is present in the atmosphere as N_2 that cannot be used by the majority of living organisms (Roswall, 1983). Moreover, nitrogen in combined forms with hydrogen or oxygen (such as NO_3^- and NH_4^+), that can be assimilated by most organisms, is highly soluble in water and can be leached by rain waters or transformed in volatile gases through a process called volatilization (Mellor et al., 1989; Viero et al., 2014).

Since nitrogen is, after water, the second most limiting nutrient for plant growth, nitrogen fertilizers are produced through synthetic methods in order to apply them in food production fields. This is possible through a process named Haber-Bosch. However, this process demands a large amount of fossil sources to produce the fertilizers. It is estimated that the production of 1 kg of NH_3 required 7.9 kWh of energy and that the production process of 1 ton of NH_3 generates an emission of 1.66 to 2.86 tons of CO_2, which represents 1.44% of the global CO_2 emissions (Soloveichik, 2019).

The world current demand of nitrogen fertilizers already surpasses 100 million of tons (FAO, 2019), being mostly intended for the fertilization of wheat, maize and rice (Ladha et al., 2016). Despite the application of large amounts of synthetic nitrogen in production fields, it is estimated that only 25 to 50% of nitrogen is efficiently assimilated by the plants for biomass production (Cassman and Dobermann, 2022; Ladha et al., 2016). The low efficiency in the use of nitrogen in fields is also associated with environmental problems. The nitrate maintenance in the soil becomes a source of leaching by rain and irrigation waters, ending in ground and surface water bodies and causing eutrophication processes. In some production fields, the losses of nitrate can reach more than 30% (Hussain et al., 2019). The annual global nitrogen flows into croplands were estimated by (Liu et al., 2010) as approximatively 140 Tg.N. In these systems, nitrogen fertilizers contributed for almost 50% of the total nitrogen flow. In the same study, nitrate leaching corresponded to 23 Tg.N/year. Besides the losses caused by leaching, the nitrate and the ammonium not used by plants may also become a substrate for the synthesis of greenhouse gases, *i.e*. N_2O and NO (Menegat et al., 2022; Viero et al., 2014). These

processes can be particularly accelerated under edapho-climatic conditions that do not favor biomass accumulation in the soil, for example high temperatures and frequent rainfalls, since the mineralization of these materials also increases the soluble nitrogen content in the soil (Raich et al., 2006).

A group of bacteria, the nitrogen-fixing bacteria, is able to use the atmospheric nitrogen due to the presence of an enzyme, the nitrogenase, which reduces the triple bond in N_2 to form ammonia, allowing the assimilation of this nutrient. Among these organisms, rhizobia are bacteria able to fix atmospheric nitrogen into ammonia in a mutualistic interaction with legumes plants. These interactions imply complex morphological changes in both partners, resulting in the formation of new organs on the roots of the plant partner, the nodules (Masson-Boivin et al., 2009). Nodulation in legumes is obtained through a highly specific interaction between the plant host and the rhizobia, which includes the recognition of both partners, infection and colonization of the roots by bacteria, formation of the nodules by the plant and, finally, nitrogen fixation. In nature, this group of bacteria was shown to be able to form nitrogen-fixing symbiosis with more than 16,000 species of legumes (Fabaceae), spread in more than 650 genera (Downie, 2014; Griesmann et al., 2018) and the non-legume *Parasponia* spp., which is able to develop nitrogen-fixing symbiosis with a wide range of rhizobia even being classified in the clade of the Rosales (Op den Camp et al., 2012).

These interactions have important agronomical consequences today. The use of legume plants combined with other plant species in crop fields has been shown to be an interesting strategy to improve nitrogen contents in the soil and consequently, decrease the need for nitrogen fertilizers in food production fields. The intercropping of maize and faba bean inoculated with rhizobia was responsible for enhancing the productivity and reducing nitrogen losses in a desert soil (Mei et al., 2021). The addition of biomass from *Gliciridia sepium* (Jacq.) in maize fields was shown to increase by 50% the nitrogen content of maize plants (Sena et al., 2020). The same was observed in wheat fields. The co-cultivation of wheat and cowpea (*Vigna unguiculata* (L.)Walp.) resulted in increased nitrogen accumulation by both plants, improved nitrogen use efficiency and higher wheat grain yield (Galindo et al., 2022).

1.2. Establishment of symbiosis between rhizobia and legumes

The symbiotic interaction between rhizobia and legumes starts with the entry of the bacterial partner into root tissues. Rhizobia can infect the root via two major strategies: either by the

attachment to root hair and the formation of the so-called infection threads (IT) or by crack entry. In the first case, rhizobia attach to the tips of the root hairs, which deform and curl around one or two bacteriaupon the production of Nod Factors by rhizobia (see more details in the sections 2 and 3 of this thesis introduction) (D'Haeze & Holsters, 2004; Downie, 2010). As a consequence of the root hair curling, bacteria are trapped in a confined space (Sahlman, 1963), then grow and divide, leading to the formation of a microcolony, which develops within an infection chamber (Gage, 2004). The high amount of bacteria entrapped within the infection chambers generates a high concentration of Nod factors, which may induce the initiation of the infection threads (Catoira et al., 2000; Walker and Downie, 2000). Then, this structure invaginates and forms the infection threads that will elongate and progress across the root cell layers until cortical cells, where bacteria will be released from the infection thread into the intracellular space (Gage, 2002). Inside the infection threads, rhizobia may be surrounded by an exopolysaccharide capsule, which were suggested to be important for the movement of the bacteria within this compartment (Rathbun et al., 2002; Roth and Stacey, 1989).

In the second case, rhizobia enter the roots by crack entry at the junctions of lateral roots or wounds. At this step, bacteria infect the root only intercellularly, but can trigger plant cell death and form infection pockets, from which they either form of infection threads that are then released into cortical cells or directly invade cortical cells (Capoen et al., 2010; Madsen et al., 2010). In certain *Aeschynomene* species, bacteria enter the root *via* crack entry, but are released into a single cortical cell, which begins to divide, consequently forming a nodule primordium.

In parallel to root entry, nodule organogenesis is initiated through the activation of a signaling cascade (see sections 2 and 3 on this introduction for more details), generally (but not always) upon the perception of the Nod Factors produced by bacteria. Consequently, two spatially distinct programs are activated: an epidermal program that allows the infection of the roots by the bacteria and a pericycle /cortical cells program allowing the infection and formation of nodule primordia (Lebedeva et al., 2021; Oldroyd and Downie, 2008). In general, the infection of legume roots takes place in the epidermis, while nodule primordium initiates in the pericycle and cortex cells (Liu et al., 2019).

In terms of morphology, the nodules can be separated in two types depending on the legume species: determinate and indeterminate. These two types are different regarding the position

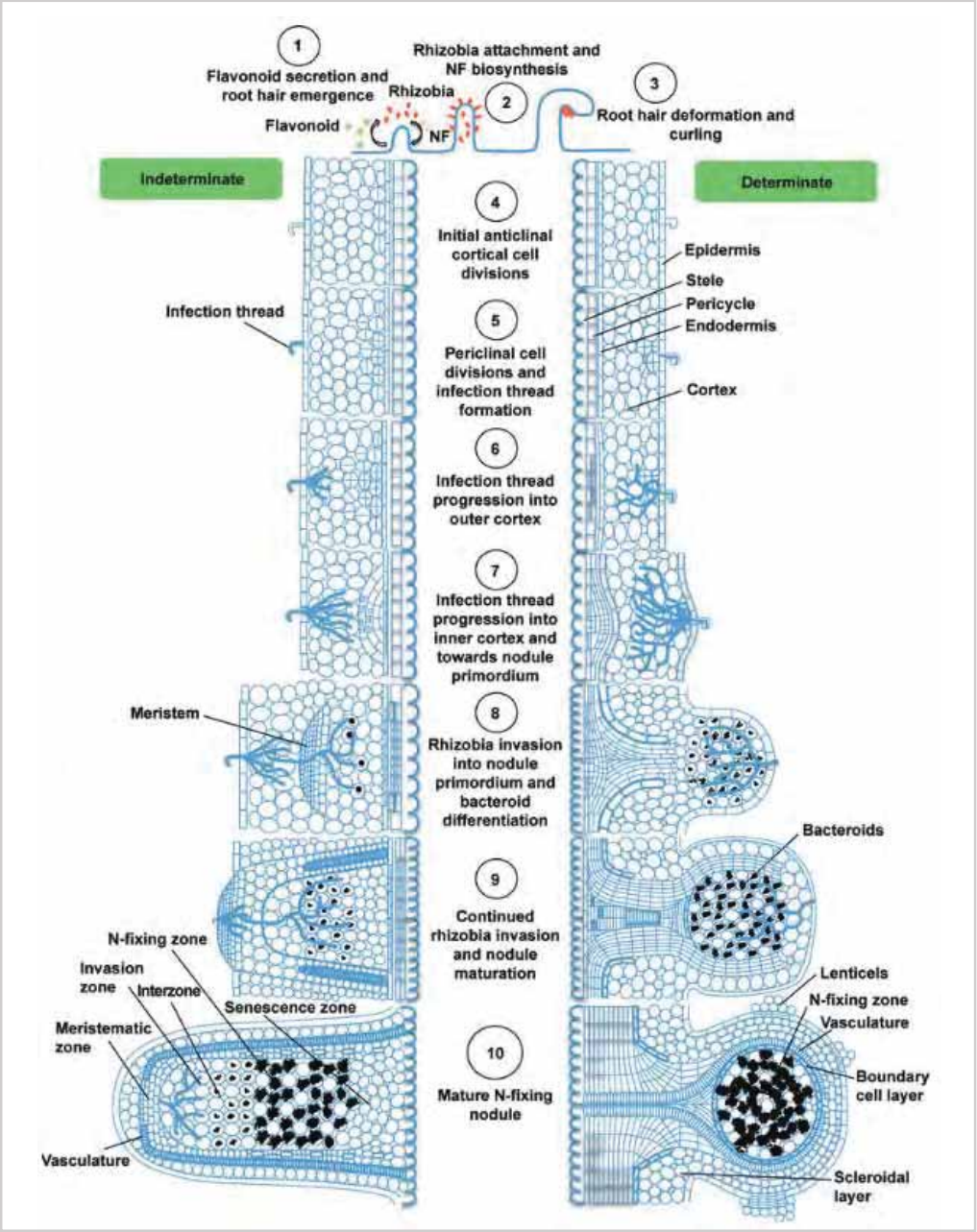


Figure 1.1. Development of indeterminate and determinate nodules. In brief, following the exudation of flavonoid compounds by the plants (1), compatible rhizobia are attracted and synthesize the so-called Nod factors (NF, 2). Both flavonoids and Nod factors determine the compatibility of the bacterial and plant species. Then, root hairs curl and form pockets where rhizobia are entrapped (3). At this point, a strong selective bottleneck acts on the rhizobial partners, since generally only one bacterium will be able to enter in each root hair and induce the formation of each individual nodule. Later, the rhizobia enter the plant host via infection thread structures that initiate in the pocket. Following this step, bacteria induce plant cell divisions in either the inner cortex or sub-epidermal cell layer, for indeterminate or determinate nodules, respectively. Then, the nodule primordium is formed by additional cell divisions of cell layers. The infection thread growth in direction to the primordium and releases the rhizobia into infection droplets called symbiosomes (4-7). Finally, bacteria differentiate into nitrogen-fixing bacteroids (8-10). In indeterminate nodules, cell divisions occur in the inner cortex, then other divisions occur in the pericycle and endodermis, leading to nodule primordium. In this case, an apical meristem develops continuously and produce new cells that can be infected by the bacteria. This means that mature indeterminate nodules contain successive zones of bacterial invasion where heterogeneous populations of nitrogen-fixing bacteria are located. On the contrary, in determinate nodules, the first cell divisions occur in the sub-epidermis in the outer cortex. Mature nodules contain relatively homogeneous nitrogen-fixing bacteroid populations, as the differentiation occurs synchronously, due to the absence of a persistent apical meristem and of the constant formation of new nodule cells caused by the presence of this plant tissue (extracted from Ferguson et al., 2010).

of the first cell divisions, nodule shape and homogeneity of the nodule cells. Indeterminate nodules are present in legumes from three subfamilies of Leguminosae, such as *Medicago sativa*, *Pisum sativum*, *Mimosa pudica* and *Vicia sativa*. In this case, cell divisions initiate in the inner cortex, followed by periclinal divisions in the endodermis and pericycle. In general, cell divisions occurring in the endodermis and pericycle allow the formation of uninfected tissues, while cell divisions at the inner cortex form the infection zone (Xiao et al., 2014). These nodules also present a persistent apical meristem, which is responsible for their typical cylindrical shape. Once formed, the nodules are not homogeneous, having distinct zones inside. In the meristem (zone I), plant cells actively divide and do not contain any bacteria. In the zone II, the bacterial cells are released from the infection thread, then bacteria differentiate into bacteroids surrounded by the so-called symbiosome membrane. In the transition between the zones II and III, an oxygen barrier is present, as well as plastids containing starch. In the zone III, nitrogen fixation itself is performed by nitrogen-fixing bacteroids. Bacterial and plant cells begin to degenerate in the zone IV and bacteria recover a saprophytic state in the zone V (Timmers et al., 2000). On the contrary, in determinate nodules, for example the nodules formed by *Glycine max*, *Lotus japonicus* and *Phaseolus vulgaris*, cell divisions occur in the middle or outer cortex, leading to spherical shaped nodules lacking a persistent nodule meristem (Saeki, 2011; Sprent & James, 2007). In these nodules, infected cells divide synchronously and mature nodules contain a relatively homogeneous population of nitrogen-fixing bacteroids (figure 1.1). In both types of nodules, released bacteria in nodule cells are surrounded by a host membrane, forming a symbiosome, and differentiate into nitrogen-fixing bacteroids (Coba de la Peña et al., 2017).

1. 3. Rhizobia are phylogenetically diverse bacteria

The term rhizobia is a generic term, which defines bacteria able to form N_2-fixing nodules on legumes. These bacteria were found in very distant phylogenetic branches among α and β-proteobacteria (figure 1.2) and designated as α- and β-rhizobia (Masson-Boivin et al., 2009). In some cases, these bacteria are interspersed in the same phylogenetic group with non-rhizobial species, including pathogens (diCenzo et al., 2019). So far, more than 180 species of rhizobia have been described. α-rhizobia are spread among 15 genera and are commonly able to interact with legumes of agricultural importance, such as soybean, peanut and chickpea, while β-rhizobia (Moulin et al., 2001) were identified in only three genera, *Cupriavidus*,

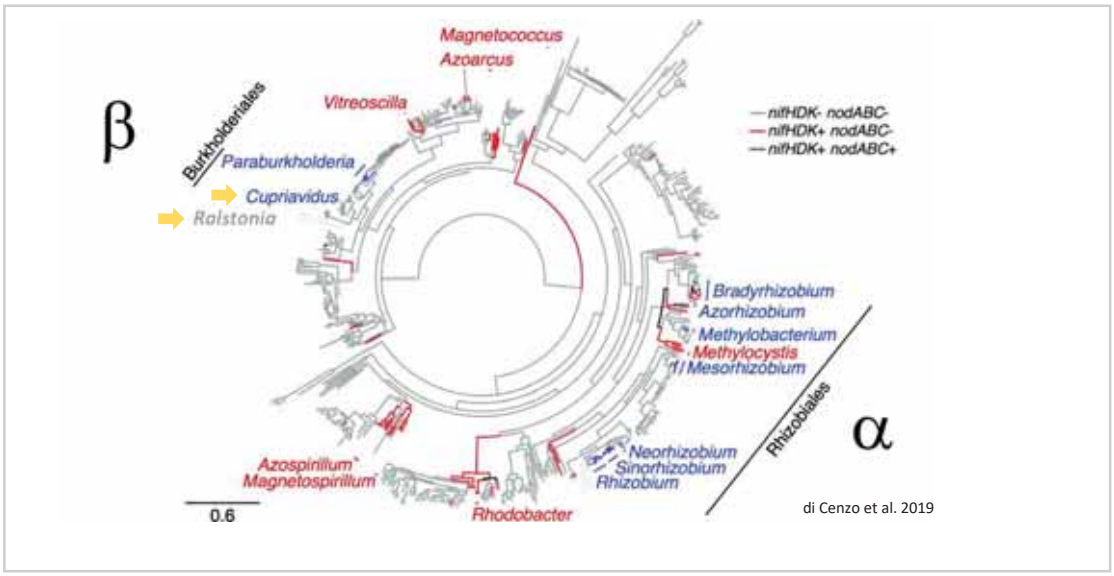


Figure 1.2. Rhizobia diversity among α and β-proteobacteria. Phylogenetic three of α and β proteobacteria constructed using 23 sequences of highly conserved proteins (Frr, NusA, RplB, RplC, RplD, RplK, RplL, RplM, RplN, RplP, RplS, RplT, RpmA, RpoB, RpsB, RpsC, RpsE, RpsI, RpsJ, RpsK, RpsM, RpsS and Tsf). In blue, bacterial species carrying genes linked to nitrogen fixation (*nifHDK*) and to nodulation (*nodABC*). In red, bacterial species only carrying genes linked to nitrogen fixation. Bacteria carrying the nodulation genes but not the nitrogen fixation genes were not identified in this work. Gray lines represent bacterial species lacking both nodulation and nitrogen fixation genes. Yellow arrows indicate the donor of the symbiotic plasmid in the evolution experiment and the recipient bacterium, *Cupriavidus* and *Ralstonia*, respectively. (Adapted from Di Cenzo et al., 2019).

Paraburkholderia and *Trinickia* (Chen et al., 2021; Tang and Capela, 2020). Besides α- and β-rhizobia, a pioneer study have described strains of *Pseudomonas* sp. nodulating the legume black locust (*Robinia pseudocacia*), which would suggest the presence of rhizobia also in the group of γ-proteobacteria (table 1.1) (Shiraishi et al., 2010). The presence of γ-rhizobia has been confirmed in fields (González et al., 2019) and tested under laboratory conditions (Bañuelos-Vazquez et al., 2019).

1. 4.Emergence and spread of rhizobia in different taxa

1.4.1.Repeated horizontal transfers of essential symbiotic genes are at the origin of rhizobium diversity

The ability to nodulate and fix nitrogen is conferred by a set of essential symbiotic genes, mainly genes related to Nod factors production (*nod* genes) and nitrogen fixation (*nif-fix* genes). Sequencing of rhizobial genomes has shown that these essential symbiotic genes are located on mobile genetic elements, more specifically on large plasmids or within genomic islands (MacLean et al., 2007) that can be transferred between bacterial species within and between genera (Andrews et al., 2018; Masson-Boivin et al., 2009; Sullivan et al., 2002). Indeed, in spite of the phylogenetic diversity of rhizobia, some symbiotic genes, such as the common *nodABC* genes, have a monophyletic origin and many of them are similar between distant species of rhizobia (Andrews et al., 2018; Bontemps et al., 2010; de Lajudie & Young, 2019; Masson-Boivin et al., 2009; Steenkamp et al., 2008).

Horizontal transfers of symbiotic genes were observed by (Sullivan et al., 1995) in a pioneer field experiment. In this work, *Lotus corniculatus* inoculated with a *Mesorhizobium loti* compatible strain was cultivated in a field devoid of rhizobia capable of nodulating this plant species. After seven years, nodules obtained from these plants were occupied by indigenous rhizobia that acquired horizontally the symbiotic island of the *M. loti* inoculated strain. Very similar results were observed by (Nandasena et al., 2006; Nandasena et al., 2007). *Bisserula pelecinus* was inoculated with *Mesorhizobium ciceri bv. Biserrulae* and cultivated in fields. Although *B. pelecinus* develops an extremely specific symbiotic interaction with *M. ciceri bv. Bisserulae*, bacteria isolated from nodules after a few years were different from the original inoculant and exhibited high competitiveness for nodulation but were suboptimal or ineffective for N_2 fixation (Nandasena et al., 2007). In Brazil, (Barcellos et al., 2007) also observed indigenous rhizobia capable of nodulating soybean after acquiring symbiotic genes

Table 1.1: Bacterium genera containing rhizobia and examples of rhizobium species.

Class of proteobacteria	Genera*	Examples of rhizobial species
Alpha	*Allorhizobium*	*qilianshanense, taibaishanense, undicola*
	Aminobacter	*anthyllidis*
	Azorhizobium	*caulinodans, doebereinerae*
	Bradyrhizobium	*betae, canariense, denitrificans, elkanii, japonicum (diazoefficiens), liaoningense, retamae, yuanmingense*
	Devosia	*neptuniae*
	Mesorhizobium	*albiziae, amorphae, chacoense, ciceri, huakuii, loti, mediterraneum, plurifarium, septentrionale, temperatum, thiogangeticum, tianhanense, opportunistum*
	Methylobacterium	*nodulans*
	Microvirga	*lupini, lotononidis, ossetica, vignae, zambiensis*
	Neorhizobium	*galegae, huautlense*
	Ochrobactrum	*lupini, cytisi*
	Pararhizobium	*giardinii, herbae, helanshanense, sphaerophysae*
	Phyllobacterium	*salinisoli, sophorae, trifolii, zundukense*
	Rhizobium	*altiplani, cellulosilyticum, daejeonense, etli, freirei, gallicum, hainanense, huautlense, indigoferae, leguminosarum, leucaenae, loessense, lupinii, lusitanum, mesoamericanum, mongolense, sullae, tropici, yanlingense*
	Shinella	*kummerowiae*
	Sinorhizobium (Ensifer)	*adhaerens, americanum, alkalisoli, arboris, fredii, kostiense, kumerowiae, medicae, meliloti, mexicanus, morelense, saheli, sojae, terangae, xinjiangense*
Beta	*Cupriavidus*	*taiwanensis, necator, pinatubonensis*
	Paraburkholderia	*caribensis, diazotrophica, mimosarum, nodosa, phenoliruptrix, phymatum, piptadeniae, ribeironis, sabiae, tuberum*
	Trinickia	*symbiotica*
Gamma	*Pseudomonas***	

* Rhizobia taxonomy is based on (Mousavi et al., 2015 ; Ormeno-Orrillo et al., 2015 ; Estrada-de Los Santos et al., 2018). ** Need complementar and independent validations. Extracted from Tang and Capela (2020) and complemented with Gonzalez-Hernandez et al., 2018.

from *Bradyrhizobium japonicum* inoculated on soils. Moreover, the sequencing of a rhizobium population from rhizosphere and root nodules of bean growing in the same field indicated that symbiotic genes massively spread between closely related species within a soil population (Pérez Carrascal et al., 2016). By showing incongruences between rhizobium core and symbiosis genes, genomic and phylogenetic analyses also provide evidences of the common occurrence of horizontal transfers of symbiosis genes both within and between genera (Andrews et al., 2018). In cases where recipient and donor strains are closely related, these horizontal gene transfers may directly convert the recipient bacterium into an efficient legume symbiont. For example, the transfer of a megaplasmid from *Rhizobium galegae* to non-nodulating *Rhizobium meliloti* mutants restored their capacity of nodulating *Medicago sativa* and changed their host specificity, inducing nodulation and nitrogen fixation on *Galega orientalis*. In this same work, the transfer of symbiotic genes from *R. galegae* to *Agrobacterium radiobacter*, a bacterium from a closely related genus (Suominen et al., 2001), led to nodulating and nitrogen-fixing strains (Novikova & Safronova, 1992). Likewise, transfer of *Rhizobium tropici* symbiotic plasmid into *Ensifer adhaerens* led to nodulation and nitrogen fixation (Rogel et al., 2001). Altogether, these observations indicate that horizontal transfers of symbiotic genes have played a major role in the emergence of rhizobia and in the dispersal of the symbiotic capacities among diverse bacterial groups although successful transfers occur preferentially between phylogenetically close organisms (Pérez Carrascal et al., 2016; Remigi, 2016) and are probably rare between α- and β-proteobacteria (Andrews et al., 2018; Lemaire et al., 2015).

The analysis of 1,314 bacterial genomes within the Rhizobiales order revealed multiple gains and losses of key symbiotic (*nod, nif*) genes, well correlated with the symbiotic capacity of the strains, and is consistent with the hypothesis that the ancestor of rhizobia lacked the symbiotic gene kit (Garrido-Oter et al., 2018). The *nod* genes were probably transferred to rhizobia from an ancestor of the genus *Frankia*. Indeed, the bacterial strain Dg1 of *Candidatus Frankia datiscae* was shown to possess the canonical *nod* genes *nodA*, *nodB* and *nodC* and phylogenetic analysis of these genes evidenced that Dg1 Nod proteins were at the root of α and β-rhizobial NodABC (Persson, 2015). Additionally, a *Rhizobium leguminosarum* A34 *nodC*::Tn5 mutant had the symbiotic phenotype restored by the *nodC* gene from the Dg1 strain of *Frankia*. The presence of the *nod* genes does not seem to be a general feature of

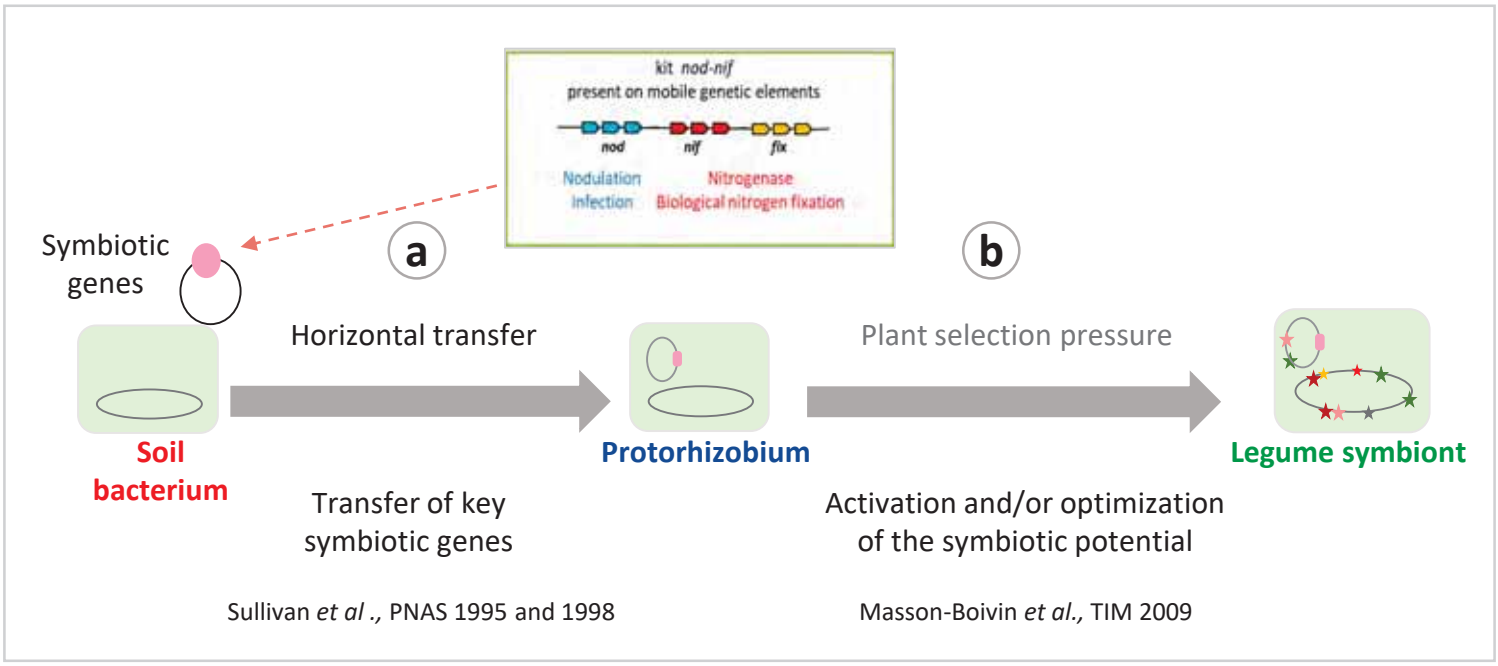


Figure 1.3. The emergence and spread of rhizobia likely occurred in a two-step evolutionary scenario. Following the acquisition of some essential symbiotic genes (a), the recipient bacterium is exposed to one or more steps of activation and/or optimization of the acquired symbiotic potential (b) in contact with the host plant. The post-transfer adaptation step likely includes the activation of some necessary biological functions or the inactivation of some deleterious functions in the recipient bacterium as well as allelic variations, gene duplications and neofunctionalization.

Frankia. Although some ancestral symbiotic *Frankia* strains were shown to possess the canonical *nod* genes (Berckx et al., 2022), some other strains do not carry them.

Noteworthy, symbiotic genes are sometimes located on mobile genetic elements that are not easily transferable, either because these elements are very large, or because they lack some conjugative functions which make them non self-mobilizable, or because their transfer involves complex regulatory mechanisms (Wardell et al. 2021). For example, very large symbiotic plasmids, possibly resulting from a long evolution and several events of acquisition and integration of new genes, may be more difficult to transfer than small mobile genetic elements carrying a less complex set of genes. This is the case, for example, of the symbiotic plasmid of *Sinorhizobium meliloti* strain 1021, that carries more than 1 Mb in addition to the symbiotic cassette (composed by 58 genes, or just 63 Kb). In this symbiotic plasmid, the other "non-symbiotic" genes seem to be important for nodulation, since strains carrying just the 58 essential symbiotic genes were less competitive for this symbiotic trait (Geddes et al., 2021). These other genes were predicted to be related to survival under low oxygen conditions and expansion of the metabolic profile (Barnett et al., 2001; Oresnik et al., 2000). Similarly, ICE replicons from *Mesorhizobium loti* can be organized in a tripartite configuration, undergoing recombination events that divide the replicon into three non-contiguous sections, but excise and integrate as one replicon (Haskett et al., 2016; Haskett et al., 2018). In this case, their transfer requires a fine tuned regulation, since the three parts have to move together and each movement is controlled by a different integrase, becoming a complex process, compared to the transfer of simple mobile genetic elements (Wardell et al., 2022).

Finally, the recruitment of *nod* genes for symbiosis is not always necessary and in some cases, rhizobia are able to nodulate leguminous plants through a *nod*-independent mechanism, however this mechanism remains to be fully elucidated. This is the case, for example, of the interaction between *Bradyrhizobium* and some *Aeschynomene* plants (Giraud et al. 2007; Teulet et al. 2019).

1.4.2. Evidence for a two-step evolutionary scenario

Although horizontal transfers have played a major role in the dissemination of rhizobial symbiotic traits, it may not be sufficient to convert directly a recipient bacterium into a legume symbiont, in particular when the recipient genome is not compatible with symbiosis. For examples, the transfer of diverse rhizobium symbiotic plasmids into *Agrobacterium*

tumefaciens or *Escherichia coli*, led to strains that were only able to form either pseudo- or non-functional nodules (Abe et al., 1998; Hirsch et al., 1984b; Nakatsukasa et al., 2008) or nodules fixing low amounts of nitrogen (Martinez et al., 1987). In line with this, suboptimal symbionts, forming non-fixing nodules, have been observed in the wild following the horizontal acquisition of a symbiosis island (Nandasena et al., 2007). Based on these observations, a two-step scenario for the evolution of rhizobia has been proposed: the evolution of a new rhizobial species requires a post-transfer adaptation step allowing the recruitment and/or the inactivation of specific functions of the recipient genome, especially if the transfer occurred between bacteria with different genomic backgrounds and/or lifestyles (figure 1.3). This optimization step likely occurred via genome remodeling under plant selection pressure (Masson-Boivin et al., 2009).

1.4.3.Many bacterial genes are required for symbiosis

Despite the importance of the *nod* and *nif-fix* genes, they are not the sole determinants of the symbiotic capacities of rhizobia. Indeed, several other genes located outside the symbiotic island/plasmids have been described as engaged in symbiosis. *In planta* transposon-sequencing (Tn-seq) approaches have suggested that 8 to 15% of the *Sinorhizobium meliloti* genes are involved or play a role during the different symbiotic stages (Flores-Tinoco et al., 2020; Pobigaylo et al., 2008). In the same way, 8% of the genes of *Rhizobium leguminosarum* were shown to be required for the competitive ability of this bacterium to nodulate pea and fix nitrogen. Among these genes, only 27 were annotated as *nif* and *fix*. Other genes were predicted to affect bacterial motility, cell envelope restructuring, nodulation signaling, nitrogen fixation and metabolic adaptation (Wheatley et al., 2020). Using the same method, (Salas et al., 2017) found that 2% (*ca.* 175 genes) of *S. meliloti* genes are involved in alfalfa rhizosphere colonization, another important stage directly impacting symbiosis establishment. These bacterial genes allowing the symbiosis with legumes are spread on the entire genome and affect several biological functions, such as metabolism, gene regulation and other cellular processes, which confirms that symbiosis requires a diversity of genes (see paragraph 2 of this introduction for a description of bacterial genes involved in symbiosis). Interestingly, a significant part of these genes are lineage-specific (Black et al., 2012; Tian, Zhou, et al., 2012), suggesting an extensive recruitment of local genes for symbiosis.

1.5. Particular circumstances allowing the successful actualization of symbiotic potential

Horizontal gene transfers are highly promoted in the rhizosphere and nodule environments (Bañuelos-Vazquez et al., 2019; Ling et al., 2016). Consistent with this observation, the horizontal transfers of symbiotic genes were shown to be frequent in nature (Lemaire et al., 2015; Pérez Carrascal et al., 2016). It is likely that exceptional circumstances are needed for the emergence of new rhizobia. Field experiments allowed to define these circumstances: introduction of an exotic legume with its natural compatible rhizobium to a new environment, the presence of soil bacteria adapted to the environmental conditions acting where the legume is introduced, and the absence of competing rhizobia (Chen, James, et al., 2003; Suominen et al., 2001; Zgadzaj, 2015).

1.6. Need for predisposing functions to achieve symbiosis with legumes?

Almost all the genera of proteobacteria existed and had already diverged (Marin et al., 2017; Turner & Young, 2000) when the legume plants emerged on earth (*i.e.* 60-100 million years ago) (Sprent and James, 2007; Werner et al., 2014). However, nitrogen fixation traits were only observed in 15 α and 3 β-proteobacterium genera (Tang & Capela, 2020) and putatively in one genus of γ-proteobacteria (González et al., 2019). This may suggest that just a subset of proteobacteria were predisposed for symbiosis.

Using comparative genomics, (Garrido-Oter et al., 2018) analyzed the genome sequences of 1,314 bacterial strains from the Rhizobiales order. Strains isolated in this work were from taxonomically distant hosts, including non-legumes, soil, insects and nematodes. Despite their variable origins, there was no whole-genome level signature of adaptation in strains interacting with legumes. All strains were separated by taxonomic lineages but neither by their ability to nodulate nor by the legume plants they colonize, nor by their origin (soil and root isolates). However, analysis of Rhizobiales communities in soil, root, and rhizosphere samples suggests that the majority of soil-borne Rhizobiales are also adapted to the root environment. Rhizosphere pre-adaptation of legume symbionts was confirmed by the study of (Salas et al., 2017) showing that genes required for alfalfa rhizosphere colonization in *S. meliloti* are conserved in distant members of Rhizobiales. The rhizosphere is a very particular environment, imposing different conditions to bacterial growth compared to bulk soil. Due to the presence of plant root exudates and water uptake, the rhizosphere is known as an environment of high osmolarity. Interestingly, in bacteria, high osmolarity can induce the

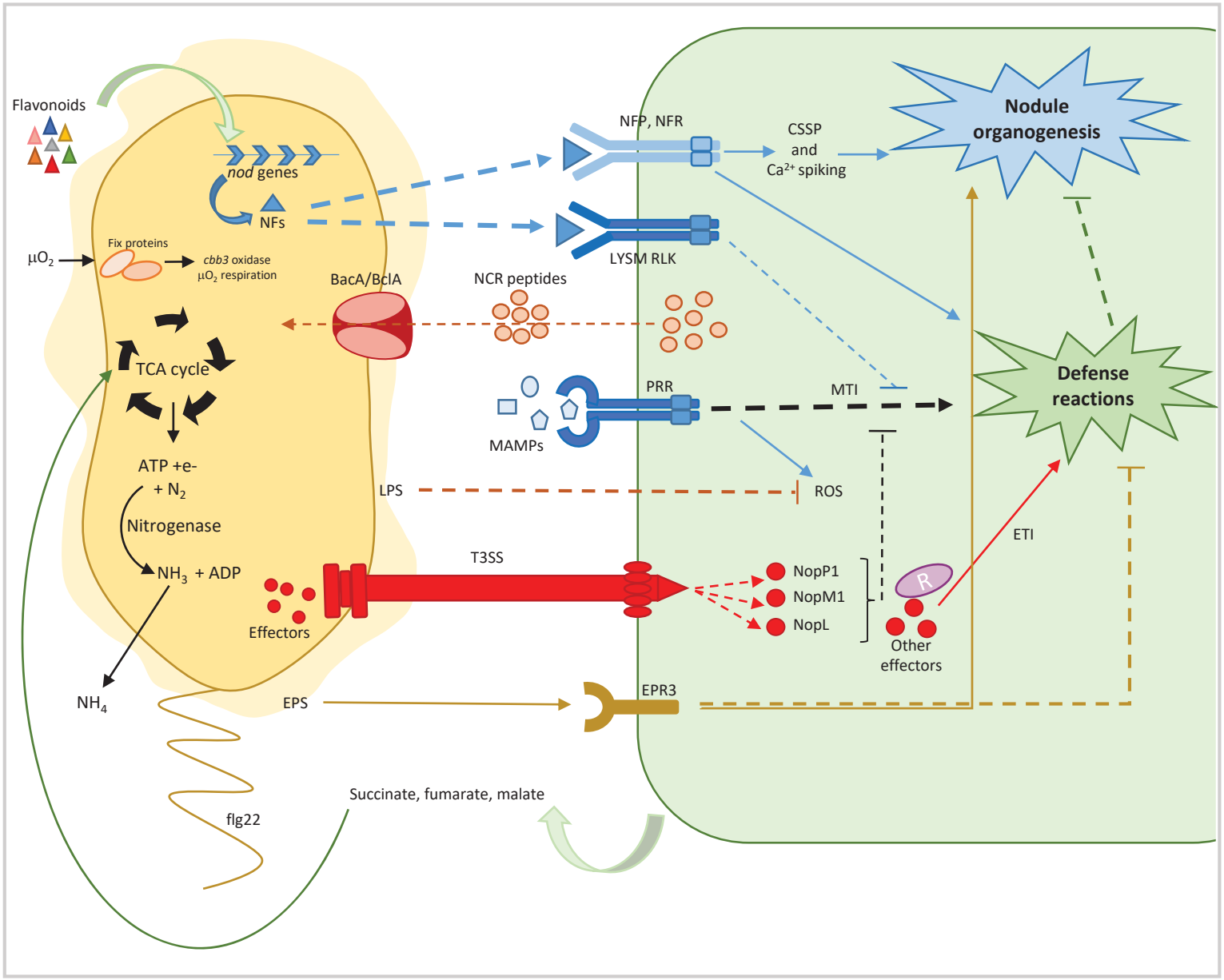


Figure 1.4. The complex molecular dialog between legume plants and rhizobia during the nitrogen-fixing symbiosis. Plants are able to exude flavonoids in the rhizosphere, which are perceived by the NodD proteins of rhizobia, leading to the induction of *nod* genes. This induction leads to the production of Nod Factors (NFs) that are recognized by specific plant receptors (symbiotic LysM-RLKs receptors NFR/NFP). The recognition of NFs has two consequences: I. the repression of the MAMP triggered immunity (MTI) of the plants, inhibiting plant defense reactions; II. The induction of the common symbiotic signaling pathway (CSSP) and calcium spiking, resulting in nodule organogenesis if both partners are compatible. Otherwise, plants are able to perceive MAMPs via PRRs that trigger MTI, resulting in calcium influx, reactive oxygen species (ROS) production and activation of an immune signaling. In some cases, rhizobia can evade plant defense responses via the production of modified MAMPs. For example, the very divergent rhizobial flg22 from *Sinorhizobium meliloti* determine the specificity of the defense reactions induced by this MAMP. Additionally, rhizobia can cross plant defenses through NF-mediated suppression of the MTI by unknown mechanisms after their recognition by LysM-RLKs (lysin motif receptor-like kinases). This recognition and suppression of plant immunity are independent of the symbiotic pathway and the recognition of NFs by the NFP/NFR receptors. Additionally, exopolysaccharides (EPS) produced by the bacteria can inhibit plant defense reactions through extracellular calcium chelation and bacterial lipopolysaccharides (LPS) are able to inhibit ROS production. Besides that, EPS also induce nodule organogenesis and infection *via* the recognition by the LysM receptor kinase EPR3. In particular rhizobia, the establishment of the symbiosis is assisted by T3 or T4 secretion systems. In these cases, symbiotic effectors (for example, Nop proteins) are injected into the host cytoplasm. Some Nop proteins (such as NopP1, NopL and NopM1) suppress plant defense responses resulting from the inactivation of MTI. On the contrary, if the bacterial effectors are recognized by NBS-LRR (nucleotide-binding site/ leucine-rich repeat) R proteins, the plant effector triggered immunity (ETI) is turned on, resulting in defense reactions. Once released from the infection threads in the nodule cells, bacteria differentiate into bacteroids able to fix atmospheric nitrogen into ammonia, a process triggered by NCR peptides (in IRLC and Dalbergioid clades), which require the bacterial BacA or BclA transporters. Since the functioning of the nitrogenase enzyme within the nodules imposes microoxic conditions to bacteria, the high oxygen affinity cytochrome cbb3 oxidase allows the respiration of bacteroids. During the interaction, carbon sources are provided by the plants as C4 dicarboxylic acids (succinate, fumarate, malate). The determinants of symbiosis presented in this figure were identified in different plant and bacterial species. Abbreviations: NFs: Nod Factors; NFP: Nod Factor perception; CSSP: common symbiotic signaling pathway; LysM RLKs: lysin motif receptor-like kinase; MAMPs: microbe-associated molecular patterns; MTI: MAMP-triggered immunity; PRR: pattern recognition receptor; LPS: lipopolysaccharide; ROS: reactive oxygen species; T3SS: Type III Secretion System; Nop: nodulation outer protein; R: resistance protein; ETI: effector-triggered immunity; EPS: exopolysaccharide; TCA cycle: Tricarboxylic acid cycle; NCR: nodule cysteine-rich. (Adapted from Gourion et al., 2014; Teulet et al., 2019; Tang and Capela, 2020).

biosynthesis of exoplysaccharides (Miller-Williams et al., 2006), molecules playing a role in plant-bacteria recognition, more specifically in the formation and development of infection threads during the nodule organogenesis (Kawaharada et al., 2015). Besides that, due to the presence of organic acid compounds in the root exudates as well as protons related to the maintenance of the net charge across the root membrane (Jones, Farrar, et al., 2003; Moreira and Siqueira, 2006), the rhizosphere has a pH decreased by 2 units compared to the neighboring soil (Faget et al., 2013). Acidic conditions are also found in the symbiosomes were bacteria are accommodated once within the nodules (Pierre et al., 2013). Given the high acidity of the symbiosome, a good tolerance to low pH probably facilitates the evolution of new rhizobia.

In another study, the comparison of 157 bacterial genomes from the genus *Sinorhizobium,* including symbiotic and non-symbiotic strains, revealed a set of 231 core genes of the symbiotic clade that was absent in non-symbiotic strains. In accordance with genomic data, a characterization of some representative strains allowed the identification of several phenotypes exclusives of symbiotic strains, including better tolerance to low pH and heat stress, while non-symbiotic strains showed a broader metabolic capacity (Fagorzi et al., 2020). Altogether, these results suggest that preexisting genetic features, including rhizosphere adaptation and stress resistance capacities, may have facilitated the emergence of mutualistic legume symbionts.

2. Rhizobial genes involved in symbiosis

2.1. Nodulation

Nodulation in legumes occurs via the activation by the microsymbiont of the so-called common symbiosis signaling pathway (CSSP, see section 3 of this introduction), followed by a downstream calcium signaling cascade that activates specific transcription factors responsible for inducing root nodule symbiosis (Gough and Cullimore, 2011; Oldroyd, 2013) (figure 1.4).

The CSSP pathway is induced by most rhizobia as a consequence of the biosynthesis of Nod Factors (NF) (D'Haeze and Holsters, 2002). The biosynthesis of these molecules is carried out by enzymes encoded by the rhizobial nodulation *nod* genes, whose expression is induced by flavonoids exuded by the plant host. The induction of the expression of the *nod* genes is controlled by NodD transcription regulators that recognize conserved DNA motifs, called *nod*-boxes. As NodD acts both as a receptor of the plant signals and a transcriptional regulator, if

Table 1.2. Presence of NF biosynthesis genes in various rhizobia.

Genes	Ac	Bd	Bs	Ml	Re	Rl	Sm	Sf	Ct	Pp
nodA	■	■		■	■	■	■	■	■	■
nodB	■	■		■	■	■	■	■	■	■
nodC	■	■		■	■	■	■	■	■	■
nodE				■		■	■			
nodF				■		■	■			
nodG							■			
nodH							■		■	■
nodI	■	■		■	■	■	■	■	■	■
nodJ	■	■		■	■	■	■	■	■	■
nodL				■	■		■			
nodM		■					■			
nodN		■		■	■	■	■	■		
nodO						■				
nodP							■	■		
nodQ							■	■	■	
nodS	■	■		■	■			■	■	■
nodT	■			■	■	■		■		
nodU	■	■						■	■	■
nodV		■								
nodW	■	■			■			■		
nodY		■								
nodZ	■	■		■	■			■		
noeC	■									
noeE								■		
noeI		■						■		
noeK				■						
noeL				■						
noeM									■	■
nolA		■								
nolE					■					
nolF						■	■			
nolG						■	■			
nolK	■			■						
nolL				■				■		
nolO		■		■				■		
nolP					■					
nolT				■						
nolU		■		■				■		
nolV		■		■						
nolW				■						
nolX				■						

Grey boxes mark the presence of the gene. Ac, *Azorhizobium caulinodans* ORS571. Bd, *Bradirhizobium diazoefficiens* USDA110. Bs, *Bradyrhizobium* sp. ORS278. Ml, *Mesorhizobium loti* MAF303099. Re, *Rhizobium etli* CFN42. Rl, *Rhizobium leguminosarum* bv. viciae 3841. Sm, *Sinorhizobium meliloti* 1021. Sf, *Sinorhizobium fredii* NGR234. Ct, C*upriavidus taiwanensis* LMG19424. Pp, *Paraburkholderia phymatum* STM815. Adapted and updated from Tang and Capela, 2020; Masson-Boivin et al., 2009 and Black et al., 2012 and completed from MaGe (http://www.genoscope.cns.fr/agc/microscope/home/index.php).

a compatible interaction between NodD and the flavonoids of the plant host is established, this interaction leads to the activation of *nod* genes (Schlaman et al., 1992). Nod Factor molecules are composed of a backbone consisting of an oligomer of chitin acylated at the non-reducing terminal residue. The common backbone of the NFs is synthesized by the NodABC proteins, while decorations of these molecules, linked to the recognition with compatible partners, are produced by *nod*, *nol* or *noe* specific genes. These substituents may be methyl (*nodS*), carbamoyl (*nolO* and *nodU*), acetyl, (*nodL* and *noeT*), sulfate (*nodH*, *nodPQ*), fucose (*nodZ*) and others (table 1.2) (Downie, 2010; Österman et al., 2014). In addition to the different decorations of the molecules, NFs present different length of the backbone, as well as different fatty acyl groups (D'Haeze and Holsters, 2002). Consequently, the production of one or several different NFs by each rhizobial strain determine its host range (D'Haeze and Holsters, 2002; Perret et al., 2000).

However, the biosynthesis of NFs is not universal in rhizobia. Some *Bradyrhizobium* strains lacking nodulation genes can nodulate *Aeschynomene* species, using a NFs-independent strategy (Chaintreuil et al., 2018; Giraud et al., 2007; Okazaki et al., 2016). Even if the molecular basis of this alternative strategy is not yet elucidated, some strong evidence indicate that key proteins of the CSSP, linked to the NF-dependent nodulation process, are also necessary in the NF-independent nodulation process (Fabre et al., 2015; Gully et al., 2018). In some NF-independent *Bradyrhizobia*, nodulation is dependent on a Type III Secretion System (T3SS) (Teulet et al., 2019). Interestingly, in the rhizobium *Bradyrhizobium elkanii* carrying *nod* genes, the T3SS was shown to activate nodulation on *Glycine max* in the absence of NFs (Okazaki et al., 2013). T3SS have been found in many but not all rhizobia, in particular they are present in *Sinorhizobium*, *Mesorhizobium*, *Bradyrhizobium* and *Cupriavidus* genera (Amadou et al., 2008; Hubber et al., 2004; Krause et al., 2002; Okazaki et al., 2016; Okazaki et al., 2009; Perret et al., 2000; Sánchez et al., 2009; Viprey et al., 1998). T3SS was shown to play both positive and negative roles in symbiosis. For example, mutations affecting *nolXWBTUV* genes that encode the T3SS in the promiscuous *Rhizobium* sp. NGR234 reduced the nodulation efficiency of this strain in many host legumes but it has neutral or negative effects on other hosts (Marie et al., 2003; Viprey et al., 1998).

In addition to T3SS, other secretion systems such as T4SS and T6SS have been shown to be important for nodulation, although their role is less understood. Mutants of *Mesorhizobium loti* for the T4SS showed a better nodulation of *Leucaena leucocephala* but are delayed in

nodulation and less competitive than a wild-type strain during the interaction with *Lotus corniculatus* (Hubber et al., 2004). In the same way, a T6SS mutant of *Bradyrhizobium* sp. (LmicA16) induced less nodules and is less competitive than the wild-type strain on *Lupinus* (Tighilt et al., 2021) Finally, the importance of the T6SS was also shown in the interaction between the *Paraburkholderia phymatum* strain STM815 and common bean. No symbiotic phenotype was observed in T6SS mutants but the expression of one T6SS of the bacterial strain (T6SS-b) was activated by the presence of germinated seeds of the host plant *Phaseolus vulgaris* and induced in *P. vulgaris* and *Mimosa pudica* nodules, suggesting a role of this system during the early and late stages of the symbiotic interaction (Hug et al., 2021).

2.2. Root infection

Once the nodules are formed, infection steps of root and nodule tissues are influenced by the presence of surface polysaccharides (*i.e*. exopolysaccharides – EPS, lipopolysaccharides – LPS or even capsular polysaccharides – KPS) (Downie, 2010; Fraysse et al., 2003; López-Baena et al., 2016). These molecules were predicted to physically protect rhizobia against several biotic and abiotic stresses or to act as signaling molecules preventing plant defense responses during the symbiotic process (Arnold et al., 2018; D'Haeze and Holsters, 2004; Geddes et al., 2014; Jaszek et al., 2014). The nature of surface polysaccharides required for symbiosis establishment vary between rhizobia. Among the molecules influencing nodule cell infection, exopolysaccharides EPSI and EPSII are particularly important for the elongation of infection threads in *Sinorhizobium, Rhizobium and Mesorhizobium* (Battisti et al., 1992; Cheng and Walker, 1998b), but are absent in the genera *Bradyrhizobium*, *Azorhizobium* and β-rhizobia (Black et al., 2012). Moreover, EPSI was shown to be unnecessary in *Sinorhizobium fredii*. In this case, other surface polysaccharides, the KPS, were shown to play a major role during infection thread formation and elongation (Kim et al., 1989). In *Rhizobium leguminosarum*, other EPS molecules synthesized by the products of the species-specific *pss* genes play an important role during the infection process (Black et al., 2012; van Workum et al., 1998).

EPS molecules also contribute to the specificity of the interaction between rhizobia and leguminous species. The effects and efficiency of the rhizobial EPS on formation and elongation of the infection threads and the infection of the plant roots are dependent on the host genetics. Indeed, the recognition of *M. loti* EPS by specific plant receptors has been recently elucidated in *Lotus* (see section 3 of this introduction) (Kawaharada et al., 2015;

Kawaharada et al., 2017). In some cases, EPS from one rhizobium species are able to restore the infection abilities of mutants from another rhizobium species while in other cases not. For example, *Rhizobium leguminosarum* bv. *viciae pssD* mutants had their infection abilities restored by the presence of EPS from *Rhizobium tropici* and *Rhizobium etli*, but not by EPS molecules from *Sinorhizobium* or *Bradyrhizobium diazoefficiens* (van Workum et al., 1998).

Abortive infection threads were also described in mutants unable to synthesize correctly LPS (Broughton et al., 2006; Campbell et al., 2002; Noel et al., 2000), and in mutants unable to synthesize or export cyclic β-glucans (Bhagwat et al., 1996; Bhagwat et al., 1992; Dylan et al., 1990), these latter molecules are neutral cyclic homopolymers of β-linked glucose residues often substituted by phosphoglycerol, phosphocolin or succinyl groups (Breedveld and Miller, 1994).

2.3. Survival within nodules

Once within the nodules, rhizobia have to survive under very specific conditions, including a microoxic environment, highly acidic conditions, high turgor pressure, specific nutrient availability and in some cases the presence of antimicrobial peptides limiting the proliferation of bacteria in nodules (figure 1.4).

A characteristic that distinguishes the nodule environment is the establishment of microaerobic conditions, with 10 to 20 nM of available oxygen. The establishment of a microoxic environment in the nodule is essential since the nitrogenase enzyme contain iron-sulfur clusters highly susceptible to the oxygen (Imlay, 2013; Preisig et al., 1996; Rutten and Poole, 2019). For this reason, the low levels of oxygen in the nodule cortex are maintained by a diffusion barrier. This creates a paradox, since rhizobia are strictly aerobic. High-affinity oxidases of *cbb3*-type acting at the end of the respiratory electron transport chain of rhizobia allows bacterial respiration in these conditions. The *cbb*3 cytochrome oxidases are encoded by the operon *ccoNOPQ* also called *fixNOQP*. They are composed of three subunits: the subunit I is encoded by *ccoN*, while the subunits II and III are encoded by *ccoO* and *ccoP*, respectively and anchored into the bacterial membrane. The subunits II and III are responsible for transferring electrons from cytochrome c2 to the subunit I (Cosseau and Batut, 2004).

Additionally to the survival to microoxic conditions, once bacteria reach the nodule cells, specific metabolic capacities are required to adapt to this environment and assimilate available carbon sources. In general, transcriptomics have revealed a downregulation of

biosynthetic pathways and an upregulation of specific metabolic pathways in bacteroids. In particular, genes related to the catabolism of C4-dicarboxylates, the main carbon source provided by the plant, formate and amino acids were shown to be upregulated in bacteroids (R. T. Green et al., 2019). By contrast, the downregulation of biosynthetic pathways is thought to be a consequence of nutrient provision by the plant hosts (Ledermann et al., 2021; Van den Bergh et al., 2018), specially amino acids. Indeed, after the differentiation of bacterial cells into bacteroids, a requirement for plant-derived amino acids by bacteria has been described. For example, transcriptomics have revealed that bacteroids present a downregulation of genes coding for enzymes related to the biosynthesis of branched-chain amino acids. Consequently, the supply of those amino acids by the plants creates a dependence of bacteroids on the host, making the branched-chain amino acid uptake a critical control point for regulation for an efficient mutualism (Prell et al., 2009). The loss of the ability to biosynthesize amino acids has been called symbiotic auxotrophy, because bacteria tend to become auxotrophic only in symbiosis with plant hosts (Udvardi and Poole, 2013). Symbiotic auxotrophy is thought to be also at the root of bacterial differentiation, since amino acid starvation may also induce the decrease in growth rate and failure to fully develop (Udvardi and Poole, 2013). However, the amino acids provided by the host may vary depending on the plant species interacting with the rhizobia. For example, in alfalfa, arginine auxotrophs of *Sinorhizobium meliloti* generated by Tn*5* mutagenesis developed an efficient symbiosis, showing a Fix^+ phenotype (Kumar et al., 2003). The same has been shown for glutamine, arginine and histidine auxotrophs (Dunn, 2015). In other cases where some amino acids are not provided by the plant, the genes linked to the biosynthesis of these amino acids are required for symbiosis to work successfully (Peltzer et al., 2008). For example, the expression of the cysteine biosynthesis pathway is upregulated in bacteroids. This probably occurs because the desulfurization of this amino acid to L-alanine and sulfur have been described as required to the production of iron-sulfur clusters of the nitrogenase and consequently the assembly of this enzyme (Ledermann et al., 2021; Zheng and Dean, 1994).

Additionally, due to the microoxic conditions within nodules, the production of energy and reduced electron carriers via the TCA cycle may not be favored since the oxidation of the electron carriers is restricted. In these cases, the accumulation of electron carrier molecules may inhibit enzymes important for the respiratory chain. In *Bradyrhizobium* bacteroids exposed to microoxic conditions very similar to those present in soybean nodules, increased

levels of NADH and NADPH caused a decrease in the activity of the respiratory-chain enzyme 2-oxoglutarate dehydrogenase (Salminen & Streeter, 1990). Possibly, a consequent arrest or deceleration of the TCA cycle may induce the bacteroids to assemble carbon into polymers of polyhydroxybutyrate (PHB), glycogen or some lipids. The accumulation of carbon in the different molecules is dependent on both bacteria and legume species, as well as the symbiotic phase (Ledermann et al., 2021). In pea, PHB was shown to be present in infection threads and broken in bacteroids, suggesting that this molecule is degraded to fuel bacteroid differentiation. Additionally, the presence of starch in the zones II and III of nodules formed on this same plant species suggests that glycogen is an important carbon source during bacteroid differentiation in pea (Lodwig et al., 2005). Besides that, the accumulation of carbon in molecules like PHB may have an adaptive role and support bacterial survival under stress and starvation conditions, for example inside senescing nodules or outside the host (Ratcliff et al., 2008). Interestingly, trade-offs between nitrogen fixation and accumulation of PHB have been described in PHB (-) mutants of *Rhizobium etli* able to fix significantly more nitrogen than the wild type strain (Cermola et al., 2000), indicating that rhizobia can divert cellular resources from nitrogen fixation to PHB biosynthesis.

Finally, legumes from the Inverted Repeat-Lacking Clade (IRLC), as well as *Aeschynomene* spp (Dalbergioid clade) synthesize the so-called Nodule-specific Cysteine Rich (NCR) peptides (Czernic et al., 2015; Mergaert et al., 2003). These peptides have antimicrobial effects (see part 3.3 of this introduction for more details) and the survival of bacterial to these NCR requires the presence of a specific bacterial peptide transporters homologous to the SbmA transporter of *E. coli*, named BacA in IRLC-nodulating rhizobia (Haag et al., 2011) or BclA in *Aeschynomene*-nodulating rhizobia (Guefrachi et al., 2015). Indeed, BacA mutants of *Sinorhizobium meliloti* are hypersensitive to different *Medicago truncatula* NCR peptides, form non-functional nodules and bacteroids show NCR-induced loss of cytoplasmic membrane integrity (Ardissone et al., 2011; Van de Velde et al., 2010). Besides the BacA transporter, another transporter, YejABEF, as well as surface polysaccharides, LPS and EPS, were shown to contribute to the resistance to NCR peptides in *S. meliloti* (Nicoud et al., 2021).

More recently, a plasmid-encoded peptidase, named HrrP (for host range restriction peptidase) was identified in some *Sinorhizobium meliloti* strains and shown to mediate the symbiotic compatibility with some *Medicago* species. HrrP peptidases are able to cleave

Table 1.3. Presence of nitrogen fixation genes in various rhizobia.

Genes	Ac	Bd	Bs	Ml	Re	Rl	Sm	Sf	Ct	Pp
fixLJK	■	■	■	■	■	■	■	■		
ccoNOPQ	■	■	■	■	■	■	■	■	■	
ccoG, ccoI	■	■	■	■	■	■	■	■	■	
ccoH	■	■	■	■	■	■	■	■		
ccoS		■	■	■	■		■		■	
fixABCX	■	■	■	■	■	■	■	■	■	■
fixM				■	■	■	■	■		
fixR	■	■	■	■			■			
fixU	■	■	■	■			■	■		■
nifA	■	■	■	■	■	■	■	■	■	■
nifHDKENB	■	■	■	■	■	■	■	■	■	■
nifQ	■	■	■	■	■			■	■	■
nifR	■	■	■	■	■	■	■	■		
nifS	■	■	■	■	■	■	■	■		
nifT †	■	■	■	■			■	■	■	■
nifU	■	■	■	■	■	■		■		
nifW, nifZ	■	■	■	■	■			■	■	■
nifX	■	■	■	■	■		■	■	■	■

Grey boxes mark the presence of the gene. Ac, *Azorhizobium caulinodans* ORS571. Bd, *Bradirhizobium diazoefficiens* USDA110. Bs, *Bradyrhizobium* sp. ORS278. Ml, *Mesorhizobium loti* MAF303099. Re, *Rhizobium etli* CFN42. Rl, *Rhizobium leguminosarum* bv. *viciae* 3841. Sm, *Sinorhizobium meliloti* 1021. Sf, *Sinorhizobium fredii* NGR234. Ct, *Cupriavidus taiwanensis* LMG19424. Pp, *Paraburkholderia phymatum* STM815. † *nifT* is also known as *fixU* or *fixT.* Adapted and updated from Tang and Capela, 2020; Masson-Boivin et al., 2009 and Black et al., 2012 and completed from MaGe (http://www.genoscope.cns.fr/agc/microscope/home/index.php).

several NCR peptides *in vitro* and *in vivo*, leading to a premature nodule senescence and defective symbiosis while increasing the fitness of bacteria within nodules (Price et al., 2015).

2.4. Mutualistic N_2-fixation

The last step of symbiosis occur due to the presence of an enzyme, the dinitrogenase, able to reduce atmospheric N_2 into NH_4^+ (figure 1.4). This enzymatic complex is encoded by the *nif* genes in all rhizobia. Even if the number of *nif* genes depends on the rhizobium genus (table 1.3), the *nifHDK* (coding for the nitrogenase enzyme complex) and the *nifENB* (coding for FeMoco synthesis) compose the minimal set of genes essential for nitrogen fixation. Other *nif* genes may complement the set of genes acting during the nitrogen fixation step itself. For example, the *nifTXQWZSUV* are linked to FeMoco synthesis. However, compared to free-living diazotrophs, rhizobia have a reduced set of *nif* genes. In this case, missing *nif* genes may be encoded by unidentified genes or housekeeping genes paralogous to *nif* genes (e.g. *iscA*/*nifU* and *iscS*/*nifS*) or compensated by plant enzymes (Nouwen et al., 2017). The expression of the *nif* genes is controlled by NifA, a transcription regulator, which is active under very low-oxygen conditions (Bobik et al., 2006; Sciotti et al., 2003). Oxygen levels are critical during the symbiosis, since the nitrogenase enzyme is sensitive to oxygen and bacteria are strictly aerobic. Additionally, nitrogen fixation is highly demanding in energy (16 ATP molecules are needed to reduce 1 molecule of N_2) and consequently, necessitates high respiration rates. This energy demand is ensured by the nodule environment where the plant partner provides dicarboxylic acids (succinate, malate, fumarate) to fuel the TCA cycle and synthesizes high amounts of leghemoglobins, proteins capable of transporting and delivering oxygen to bacteroids while keeping free oxygen concentrations at very low levels. Finally, the transfer of electrons to nitrogenase is encoded by the *fixABCX* genes (Ledbetter et al., 2017), which are highly conserved among rhizobia (Black et al., 2012).

Besides nitrogen fixation genes *per se*, genes related to the dicarboxylate transport system (Dct) are particularly important for symbiosis in many rhizobia (Yurgel and Kahn, 2005). Mutants of these genes were shown to form well-developed nodules but failed to fix nitrogen (Ronson et al., 1984). Once in the nodules, C4-dicarboxylates are assimilated via the TCA cycle. Indeed, genes coding for succinate dehydrogenase, fumarase and malate dehydogenase were shown to be upregulated in differentiated bacteroids. Besides that, mutants in these genes develop an ineffective symbiosis and fail to fix nitrogen (Udvardi and Poole, 2013). Indeed,

reduced electron carriers (NADH and reduced Flavin adenine dinucleotide $FADH_2$) and ATP molecules produced during the functioning of the TCA cycle are directly required for nitrogen fixation (Ledermann et al., 2021).

3. Host-mediated control mechanisms

3.1. Host entry is tightly controlled by the plant

In the early stages of the interaction, the ability of rhizobia to enter the roots and form nodules is mainly dependent on the capacity of rhizobia to produce compatible Nod factors (figure 1.4), lipo-chitooligosaccharides, whose structure is specifically recognized by the plant (Oldroyd et al., 2011). Bacterial Nod factor production is induced by plant flavonoids, aromatic compounds secreted by the roots of the legume partner. Legumes are able to synthesize a wide range of isoflavonoids, but only some of them can activate the bacterial LysR-type transcriptional regulation NodD proteins controlling nodulation genes (*nod* genes) (Boivin and Lepetit, 2020 ; Dong and Song, 2020; Peck et al., 2006). The amount and the chemical structure of the flavonoids exuded in the rhizosphere can determine the efficiency of response of rhizobia to these molecules, limiting the range of possible bacterial partners. For example, the *nod* genes of *Cupriavidus taiwanensis*, the symbiont of *Mimosa pudica*, were shown to be induced by the flavonoid luteolin (Amadou et al., 2008). In soybean, the flavonoids genistein and daidzein have a proven effect of inducing the biosynthesis of Nod factors on *Bradyrhizobium japonicum* (Kosslak et al., 1987; Pueppke et al., 1998; Pueppke and Broughton, 1999). In *Sinorhizobium meliloti*, the flavonoids naringenin, eriodictyol, dadzein and luteolin were able to stimulate the increase in the DNA binding affinity of the transcription regulator NodD1 to *nod* gene promoters, but only luteolin was able to mediate *nod* genes induction (Peck et al., 2006). In addition, flavonoids can contribute to the specificity of the legume-rhizobium interaction by negatively affecting the growth of non-target bacteria, because these molecules can act as phytoalexins (low molecular weight antimicrobial compounds) inhibiting the growth of phytopathogens or incompatible bacteria (Hassan & Mathesius, 2012) (Dong and Song, 2020; Pankhurst and Biggs, 1980). These molecules can also affect the quorum-sensing of non-target bacteria and their ability to form biofilm (Vandeputte et al., 2010). Consequently, the ability to recognize flavonoids is the first determinant of the rhizobium host range. The Nod factors produced upon activation by flavonoids are then specifically perceived by plant LysM receptor kinases NFP/NFR5/Sym37 and

LYK3/NFR1/SYM10 (Arrighi et al., 2006; Bozsoki et al., 2020; Limpens et al., 2003; Zhukov et al., 2008), which in turn activate the so-called common symbiosis signaling pathway (CSSP). The perception of the Nod factors triggers the production of Ca^{2+} spiking signals via a process that involves a set of signal transduction proteins, such as the leucine rich repeat RLK (MtDMI2 and LjSYMRK) components of a nucleoporin complex (LjNUP85, LjNIP133 and LjNENA), and nuclear membrane localized ion channels (MtDMI1 and LjCASTOR, LjPOLLUX). After the production of the Ca^{2+} spiking signal, a nuclear calcium and calmodulin-dependent protein kinase (CCaMK) (MtDMI3/LjCCaMK) interacts and phosphorylates a nuclear coiled coil protein (MtIPD3/LjCYCLOPS), which decodes the calcium spiking and activates specific transcription factors that initiate nodule organogenesis. Among them, the key transcription factor NIN acts as a hub (Liu et al., 2019) coordinating nodule development and bacterial invasion (Schiessl et al., 2019; Soyano et al., 2019). Among the early symbiotic genes activated by NIN, RPG (Rhizobium-directed polar growth) and NPL (Nodulation Pectate Lyase 1) are both required for infection thread formation (Arrighi et al., 2008; Xie et al., 2012), the transcription factors NF-YA1 and NF-YB1 play a crucial role during the nodule primordia development, nodule meristem differentiation and infection (Laporte et al., 2014), the gene *CRE1* (Cytokinin Response 1) encoding a cytokinin receptor (Inoue et al., 2001), as well as CLE and CEP genes encoding peptides that negatively and positively regulate nodulation in *M. truncatula* (Laffont et al., 2020).

Plant defense responses also participate in the control of bacterial entry. Transcriptomics have showed the induction of plant defense related genes following the inoculation of rhizobia (Andrio et al., 2013; Kouchi et al., 2004; Libault et al., 2010; Lopez-Gomez et al., 2012). Additionally, an increased production of oxygen reactive species (ROS) has been observed in legume plants after the inoculation of rhizobia and seems to be important for the infection of nodule cells, as well as for the later steps of the symbiosis (Jamet et al., 2007; Marino et al., 2011). Induction of plant immune response can rely on the recognition of the so-called MAMPs (for microbial associated molecular patterns) such as the *flg22* epitope of flagellin or microbial effectors (Gourion et al., 2015).

3.2. Host control of infection

After the recognition of Nod Factors leading to changes in plant signaling and the consequent nodule organogenesis, additional mechanisms control the efficiency of infection of the newly

formed nodules (figure 1.4). More specifically, the progression of infection depends on the production of bacterial surface polysaccharides (EPS, KPS and cyclic β-glucans), which are important in both early and late stages of the symbiotic interaction. In *Lotus japonicus,* EPS molecules, which are absolutely required for the initiation and progression of infection threads, were shown to be specifically recognized by a LysM receptor kinase, EPR3 (Exopolysaccharide receptor 3), which mediates the recognition of compatible and incompatible EPS (Kawaharada et al., 2015; Kawaharada et al., 2017). Indeed, *epr3* mutants were shown to be defective in the perception of EPS molecules and produced fewer infection threads when the plants were inoculated with *Mesorhizobium loti*. Consequently, this protein may regulate bacterial passage through the plant epidermal cell layer and infection thread progression into the underlying cortical tissues and nodule primordia. The expression of *EPR3* is dependent on the host perception of bacterial NFs (Kawaharada et al., 2015; Kawaharada et al., 2017).

3.3. Induction of growth arrest of nitrogen-fixing bacteroids

Once released in nodule cells, rhizobia located in symbiosomes start to differentiate into nitrogen fixing bacteroids (figure 1.4). This differentiation can be either terminal or partially reversible depending on the host plant, meaning that in the first case bacteroids cannot regrow from nodules (Haag and Mergaert 2019; Kondorosi et al., 2013; Mergaert et al., 2003; Montiel et al., 2016) while in the second case they can (Haag and Mergaert 2019; Kondorosi et al., 2013; Marchetti et al., 2011; Montiel et al., 2016). Terminal differentiation is characterized by particular morphology of bacteroids, either elongated (E-type bacteroids) or spherical (S-type bacteroids) and genome amplification up to 24C in *Medicago* nodules (Czernic et al., 2015; Kondorosi et al., 2013). While non-differentiated bacteroids are common in Phaseoloid or Robinioid plants, (for example *Phaseolus*, *Glycine* and *Vigna*), E-type bacteroids are common in Galegoids (for example, *Medicago*, *Vicia* and *Aeschynomene*) and S-type bacteroids are present in *Arachis*, and some *Crotalaria* and *Aeschynomene* species (Kondorosi et al., 2013). In IRLC and Dalbergioid legumes, this terminal differentiation stage is triggered by NCR peptides. The number of NCR strongly varies between legumes, only 7 were detected in *Glycyrrhiza uralensis* while more than 700 were found in *M. truncatula* (Montiel et al., 2017) and about 30 were detected in *Aeschynomene* spp (Czernic et al., 2015). The role of these peptides is not completely elucidated. In *M. truncatula*, the NCR247 peptide was

shown to interact with different bacterial cell cycle components, such as the cell division protein FtsZ (Farkas et al., 2014) and the cell cycle regulator CtrA (Penterman et al., 2014). These interactions inhibit cell divisions but allow DNA replication, which leads to the polyploidization of bacterial cells (Mergaert et al., 2006). NCR peptides also affect bacterial envelope permeability, leading to cell death *in vitro* (Van de Velde et al., 2010). However, the killing effect of NCRs *in vitro* may be due to the high concentrations of those peptides used in these assays. Consequently, the hypothesis that a moderate state of permeabilization of bacterial membranes with cationic NCR peptides may be necessary to the entry of anionic NCR peptides into the bacterial cytosol is not excluded (Kondorosi et al., 2013). This hypothesis is reinforced by the fact that in nodule cells, rhizobia are exposed to both cationic and anionic NCRs and cationic peptides interact preferentially with anionic bacterial membranes (Kondorosi et al., 2013). In addition to antimicrobial effects, NCR peptides also promote nitrogen fixation. Recently the NCR247 of *Medicago truncatula* was shown to sequester haem in *S. meliloti* bacteroids leading to an increase in transcription of iron-uptake genes, which results in an increase in iron import necessary for the activity of nitrogenase,thus improving nitrogen fixation (Sankari et al., 2022).

In symbiosomes, bacteroids rapidly stop dividing (Vasse et al., 1990). This non-growing state was confirmed by transcriptomic analyses in both determinate and indeterminate nodules producing or not NCR peptides. In particular, transcriptomics of *S. meliloti*, which undergoes terminal differentiation in *M. sativa* nodules expressing NCR peptides, evidenced a strong downregulation of the expression of most of the genes, a profile resembling the expression profile of bacteria grown in stationary phase, with the exception of nitrogen fixation genes, which are highly expressed (Capela et al., 2006). Similarly, the transcriptomic profiles of *Rhizobium etli, Rhizobium leguminosarum* and *Mesorhizobium loti*, which do not undergo terminal differentiation in *Phaseolus* and *Lotus* nodules, respectively, indicated a strong down-regulation of global gene expression in symbiosis and, in *R. etli*, an extensive overlap between genes downregulated in bacteria grown in stationary phase (R. T. Green et al., 2019; Uchiumi et al., 2004; Vercruysse et al., 2011). Altogether, these results suggest that bacteria present a growth-arrested state inside the nodules, either when exposed or not to NCR peptides, in favor of an hyper-specialization in nitrogen fixation. Additionally, experimental data combined with metabolic modeling have suggested that growth arrest of bacteroids is also explained by the metabolic constraints imposed by the nodule environment (Schulte et al., 2021). Indeed,

in nodules plants fuel bacteroids with dicarboxylic acids, whose catabolism creates a high oxygen demand. In the low-oxygen environment of the nodule, this promotes a down-regulation of the decarboxylating TCA cycle arm, which decreases ammonia assimilation into glutamate by bacteroids. Since glutamate is the transamination donor for several other amino acids, this probably contributes to the growth arrest of bacteroids. In addition, these conditions will force ammonia secretion and thus favor mutualism.

3.4. Selection of nitrogen fixation

When a rhizobium strain is able to intracellularly infect and persist within nodule cells, nitrogen fixation and nutrient exchanges take place between both partners, leading to a mutually beneficial interaction. The control of this stage is critical, since the nodulation process is costly for the plant. Moreover, as non-fixing bacteria may outcompete fixing bacteria that invest energetic resources into the process of nitrogen fixation, these control mechanisms may play a key role in the stability of mutualism over long time-scales (Friesen, 2012; Kiers and Denison, 2008). Whether plants are able to select rhizobia based on their nitrogen-fixing phenotype at the root entry point (pre-infection partner choice) and/or during the later stages of the interaction (post-infection sanctions) has thus been a long-standing subject of research.

Several studies have shown that, contrary to the ability to nodulate and enter plant tissues, plants do not discriminate at the root entry point bacterial strains able or unable to fix nitrogen (Amarger, 1981; Daubech et al., 2017; Hahn and Studer, 1986; Westhoek et al., 2017). Instead, once in the host, the legume plants exert post-infection control mechanisms to restrict the fitness of non-fixing microsymbionts (Daubech et al., 2017; Kiers et al., 2013; Laguerre et al., 2012; Oono et al., 2011; Oono et al., 2009; Simms and Taylor, 2002). Partner choice has also been observed in a few studies (Gubry-Rangin et al., 2010; Heath and Tiffin, 2009; Regus et al., 2014). However, since the nitrogen fixation process occurs in the late stages of the symbiotic interaction, it seems unlikely that direct selection on nitrogen fixation ability occurs at the root entry stage. Instead, partner choice might represent an indirect selection process for nitrogen fixation whereby the most competitive rhizobia that are also the most beneficial for the plant preferentially associate with the host plant. Strains showing a positive correlation between these two phenotypic traits might be selected in natural environments (Younginger and Friesen, 2019).

Sanctions imposed by the host were evidenced using isogenic strains of *Cupriavidus taiwanensis*, differing for the presence/absence of an active *nifH* gene. In this work, the Fix^- bacteria multiply similarly to Fix^+ bacteria in *Mimosa pudica* nodules, until 16–21 days post-infection, then they start to prematurely degenerate, consequently having a decline in their fitness compared to Fix^+ bacteria (Daubech et al., 2017). Similar results were obtained in *Acmispon strigosus* and *Lotus japonicus* nodules (Regus et al., 2017). Additionally, these sanctions occur at the nodule cell level (Daubech et al., 2017; Regus et al., 2017). However, at this point, it is still not clear whether molecular mechanisms behind these sanctions are specifically active mechanism exerted by the plant (via the activation of immune responses and/or a modulation of metabolic exchanges) or linked to a lower survival capacity of non-fixing bacteria within nodules. Possibly a combination of both may explain the premature degeneration of non-fixing symbionts. Indeed, particularly acidic conditions prevail in symbiosomes due to plant H^+-ATPases located at the peribacteroid membrane. These acidic conditions, which are buffered by the production of NH_4^+ in nitrogen-fixing symbiosomes, may prevent the survival non-fixing strains. Additionally, the allocation of oxygen (Kiers et al., 2003) and carbon sources (Jeudy et al., 2010; Westhoek et al., 2021) have been shown to be reduced in non-fixing nodules as well as the size of the nodules (Laguerre et al., 2012; Oono et al., 2011), attesting for active sanction mechanisms against inefficient symbionts.

However, these sanctions seem to be conditional, depending on the nitrogen-fixing abilities of each bacterial strain acting in the system. Using pea plants, (Westhoek et al., 2021) have observed that nodules containing intermediate-fixing strains were large and had a high carbon supply when co-inoculated with a non-fixing strain but were small and white when plants were co-inoculated with a more effective strain. This means that plants can compare a local nodule nitrogen content with a global amount of nitrogen provided by each bacterial strain infecting individually each nodule and use sanction mechanisms (such as the control of oxygen levels and allocation of carbon sources) towards less beneficial strains.

3.5. Suppression of plant defense reactions during symbiosis

Contary to other plant-microbe interactions, rhizobium-legume interactions are characterized by the absence of defense reactions and instead involve active mechanisms on both plant and bacterial side to suppress immunity (Berrabah 2019). Indeed, transient induction of plant defense-related genes, were measured in soybean, *Medicago truncatula* and *Lotus japonicus*,

few hours following the inoculation with *Bradyrhizobium japonicum*, *Mesorhizobium loti* and *Mesorhizobium japonicus*, respectively, but the expression of these genes is then rapidly suppressed (Libault et al., 2010; Lohar et al., 2006; Kouchi et al., 2004). Classical defense mechanisms developed by plants rely on the production of phenolic compounds and exudation of phytohormones, such as salicydic acid, jasmonic acid and ethylene. These molecules were shown to inhibit the development of nitrogen-fixing symbioses, indicating that they may be repressed during these interactions (Stacey et al., 2006; Penmetsa and Cook, 1997; Sun et al., 2006).

On the bacterial side, some exopolysaccharides, such as the succinoglycan produced by *Sinorhizobium meliloti* during the interaction with *Medicago truncatula*, are thought to suppress MAMP triggered immunity (MTI) in early steps of the symbiotic interaction by chelating calcium, a cellular signal involved in MTI (Aslam et al., 2018). Some rhizobia also possess a TTSS able to inject effectors within plant cells, in order to modulate MTI (Teulet 2019). In the rhizobium NGR234, the type III effector NopM inhibits plant MTI and decreased ROS production and promote nodulation on *Lablab purpureus* roots (Xin et al., 2012). In the same way, the expression of NopL in *Nicotiana benthamiana* suppressed defense responses to mosaic virus (Bartsev et al., 2003). Finally, Nod Factor molecules, vastly shown to determine bacterial-legume specificity during the initial steps of the symbiosis, have been shown to suppress plant MTI on legume and non-legume plants. This suppression seems to be independent on their recognition by the nod factor receptors NFR1 and NFR5 (Liang et al., 2013). Interestingly, rhizobial genes related to Nod Factor production were shown to be transcribed in the nitrogen fixation zone of the nodules (Roux et al., 2004). Since at this point of the interaction the nodules are already formed, one could wonder why these genes are still transcribed and whether these molecules could be related to plant defense suppression during the later steps of the interaction.

On the plant side, some genes were also shown to suppress plant immunity during the interaction with compatible rhizobia, especially during bacteroid differentiation and persistence within nodule cells. For example, nodules formed by mutants of the *RSD* gene from *Medicago truncatula* host bacteroids that senesce earlier compared to bacteroids in wild-type nodules. Nodules formed by these mutants were brownish; a marker of activation of plant defenses (Sinharoy et al., 2013). In the same way, mutants for two other plant genes, *SymCRK* (for symbiotic cysteine-rich receptor kinase) and *DNF2*, were shown to develop

necrotic nodules with a high level of phenolic compounds, molecules also related to plant immune responses. Additionally, mutants for these genes had undifferentiated bacteroids (Berrabah et al., 2014; Bourcy et al., 2013). Finally, insertions in a PLAT (α–toxin) domain encoding gene, *NPD1* (for regulator of symbiosome differentiation 1) in *Medicago truncatula* lead to ineffective and small nodules, containing non-differentiated bacteroids. Interestingly, these nodules developed normally during the early stages of the symbiosis, supporting the hypothesis that these genes are particularly important for late symbiotic steps (Pislariu et al., 2019; Trujillo et al., 2019).

On the other hand, defense reactions can be induced in case of incompatible rhizobium-legume interactions. For example, some bacterial type III effectors can induce plant defense responses if a cognate R gene is expressed in the host and triggers in a process called effector-triggered immunity (ETI). For example, in soybean, the presence of the *Rj4* allele in some ecotypes restricts the interaction with some rhizobium species (Yasuda et al., 2016). Similarly another soybean allele, *Rj2*, is able to allow or not the interaction with some *Bradyrhizobium* species (Yang et al., 2010). The product of this gene, a protein having a typical structure of a plant resistance protein called NLRs (with an interleukin receptor domain, a nucleotide binding site and leucine-rich repeat motifs), mediates the recognition of bacterial effectors and triggers plant ETI (Benezech et al., 2020). These incompatible interactions are characterized by the absence of developed infection threads. Finally, Nod factors receptors were shown to play a role in the elicitation of some plant defense responses, as shown using *Medicago truncatula nfp* mutants, which were more susceptible to the oomycete *Aphanomyces euteiches* and the fungus *Colletotrichum trifolii* (Rey et al., 2013).

4. Conversion of a plant pathogen into legume symbionts by experimental evolution

4.1. Experimental evolution as a tool to study rhizobium/legume interactions

Experimental evolution (EE) is a powerful tool to study the adaptation of organisms. It consists in propagating organisms in a controlled environment and analyzing adaptive processes in real time and in an accelerated way compared to natural evolutionary processes (Lenski et al., 1991; Manriquez et al., 2021). Coupled to the resequencing of bacterial genomes, this approach allows the identification of mutations responsible for the phenotypic changes. During the last decade, most evolution experiments have been conducted using simple systems propagating microbial populations in synthetic medium (Good et al., 2017; Gostinčar

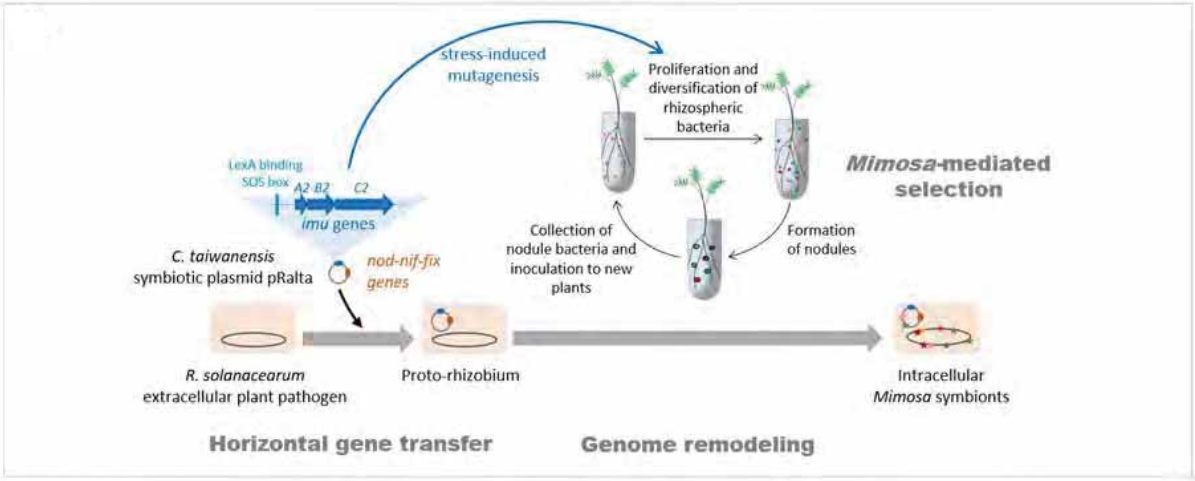


Figure 1.5. Experimental evolution of the plant pathogen *Ralstonia solanacearum* into legume symbionts. The symbiotic plasmid (pRalta) of *Cupriavidus taiwanensis*, the natural symbiont of *Mimosa pudica*, was introduced into *R. solanacearum* (Marchetti et al., 2010), generating a chimeric strain unable to nodulate *M. pudica*. After the acquisition of three nodulating variants through the inactivation of the Type III Secretion System (T3SS) in contact with the plant host, the proto-rhizobium was evolved in serial cycles of inoculation onto *M. pudica* plantlets and re-isolation of nodule bacteria in order to improve symbiotic phenotypes under selection of the host. In addition to the *nod* and *nif-fix* genes involved in Nod Factor synthesis and nitrogen fixation respectively, the pRalta plasmid carries a mutagenesis cassette (*imuA2B2C2*) that generates a transient hypermutagenesis phenomenon when the bacteria were free-living in the plant culture medium and rhizosphere. This hypermutagenesis increases the genetic diversity of the rhizospheric bacterial population among which the plant selects the best nodulating variants (extracted from Doin de Moura et al., 2020).

et al., 2021; Huang et al., 2018; Lenski et al., 1991; Lenski and Travisano, 1994) and only a small number has analysed the evolution of microorganisms interacting with their host (Cooper et al., 2020; Lescat et al., 2017; Robinson et al., 2018; Schuster et al., 2010; Tso et al., 2018). In particular, only two recent studies, in addition to the one conducted in the team, dealt with rhizobium-legume interactions (Batstone et al., 2020; Quides et al., 2021).

The objective of the first evolution experiment was to evaluate whether the hosts select for more beneficial symbionts (Batstone et al., 2020). In this experiment, 5 genotypes of *Medicago truncatula* varying in "choosiness" (*i.e.* in the relative number of nodules that they form with different bacterial strains) were co-inoculated with two rhizobial strains of *Ensifer meliloti* differing by the benefits they provide to the hosts but also by their nodulation competitiveness, one competitive and efficient nitrogen fixing strain (Em1022) and a poorly competitive and ineffective nitrogen fixer (Em1021). After five cycles of nodulation, the number of nodules formed by each strain was determined and the shoot biomass was measured. The effective symbiont, Em1022, outcompeted the less effective Em1021 strain on all hosts regardless of their choosiness. Authors concluded that partner choice was not an important selective force in the evolution of rhizobia. However, since Em1022 was initially more competitive for nodulation on all plant genotypes, this conclusion is not robust.

Besides, Em1022 clones developed stronger cooperative interactions with the hosts, on which they evolved, suggesting that local adaptation is an important force shaping the evolution of rhizobia and highlighting the fact that rhizobia are able to rapidly adapt and provide benefits to new hosts. This may have implications in the variations in the mutualistic quality of rhizobia depending on the hosts. This phenomenon may be at the root of the presence in nature of rhizobia having poor nitrogen-fixing efficiency on some hosts while being more efficient on other hosts.

The evolution of rhizobia towards increased benefits to the host plant has been tested in another evolution experiment (Quides et al., 2021). In this experiment, two strains, *Rhizobium etli* CE3 and *Ensifer fredii* NGR234, which nodulate but provide low benefits on *L. japonicus* were evolved for 10 or 15 cycles of nodulation on two genotypes of this host that vary in their regulation of nodulation. However, no evidence of evolution towards an increased nitrogen fixation was obtained during this experiment, instead a decrease in the energetic costs of the symbiosis to the host was observed. Interestingly, a stronger adaptation of rhizobia was observed when the strains were inoculated on a hypernodulating plant genotype, evidencing

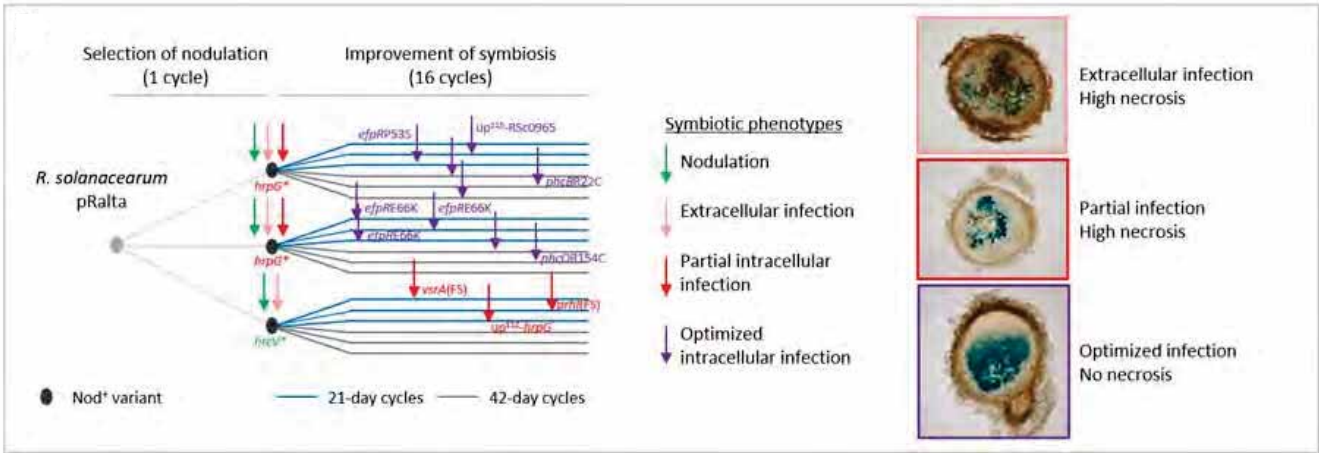


Figure 1.6. Results of the experimental evolution of the plant pathogen *Ralstonia solanacearum* into legume symbionts. 18 parallel lineages were derived from the three nodulating (Nod^+) variants via two different selection regimes (cycles of 21 or 42 days). After acquisition of the nodulation ability, bacteria progressively improved their capacity to infect nodules. Three levels of infection were observed in the evolved lines: extracellular, partially intracellular and an optimized intracellular infection. The main adaptive mutations responsible for the acquisition and improvement of symbiosis were identified. *, stop mutation. FS, frameshift. up^{112}, up^{115}, intergenic mutations located 112 bp and 115 bp upstream from the gene indicated (extracted from Doin de Moura et al., 2020).

that the number of nodules is a key factor in the adaptive evolution of rhizobia. Finally, no adaptation to the free-living steps of the interaction was observed in this study, which reinforces the importance of the plant-mediated selection of legume symbionts (Quides et al., 2021). Of note, these 2 studies used rhizobia as ancestors, and did not investigate the evolutionary origin of nitrogen-fixing legume symbionts.

4.2. Experimental design mimicking the natural evolutionary history of rhizobia

To understand the evolutionary history of rhizobia and validate a two-step scenario (part 1.4.2.), an evolution experiment was designed in the team to replay the emergence of new legume symbionts from a non-rhizobial genus under laboratory conditions (figure 1.5). The experiment started with the transfer of the symbiotic plasmid (containing the essential symbiotic genes) of the *Mimosa pudica* symbiont, *Cupriavidus taiwanensis* LMG19424, into the non-rhizobium plant pathogen *Ralstonia solanacearum* GMI1000. These two bacterial species were chosen as donor and recipient of symbiotic genes because they are neither too close to observe the post-transfer adaptation steps, nor too phylogenetically distant to increase the likelihood that these steps can occur in a reasonable time. Initially, the transfer of the symbiotic plasmid generated a chimeric strain unable to nodulate (Nod^-) *M. pudica*, showing that, in this case, the transfer of the symbiotic genes was not sufficient to convert the recipient bacterium into *M. pudica* symbionts. After several trials of inoculation of plants, three nodules were obtained from which three nodulating variants were isolated. The three nodulating variants were then further evolved under plant selection pressure in serial cycles of inoculation on *M. pudica* plantlets, isolation of nodule bacteria and reinoculation on new *M. pudica* plantlets. These cycles allowed the parallel acquisition and progressive improvement of two symbiotic traits, nodulation competitiveness and intracellular infection of nodules, in independent lineages (Marchetti et al., 2010; Marchetti et al., 2017; Marchetti et al., 2014). After 16 evolution cycles, the evolved strains of most lineages induced nodules showing features of true rhizobia nodules: a peripheral vascular system, and the presence of symbiosomes containing bacteria in nodule cells. However, mutualism was not achieved in 16 cycles and bacteria did not persist in nodule cells, even if they were able to intracellularly infect these cells (figure 1.6). This phenotype may reflect what happens in nature. Indeed, in nature rhizobia having suboptimal symbiotic abilities or inducing plant defense reactions in nodules have been observed several times (Gehlot et al., 2013; Gossmann et al., 2012;

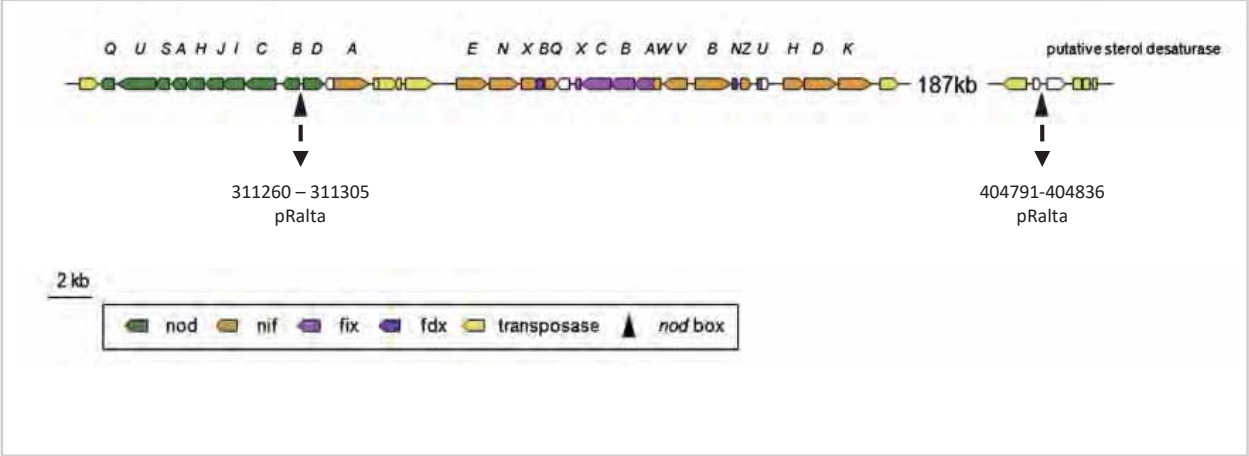


Figure 1.7. Essential symbiotic genes from *Cupriavidus taiwanensis*. Genes are colored according to their name: bacterial *nod* genes are represented in green, *nif* genes in orange, *fix* genes in purple and *fdx* genes in black purple. Transposases are represented in yellow. Black triangles indicate the location of potential nod-boxes (extracted from Amadou et al., 2008).

Nandasena et al., 2006; Nandasena et al., 2007). Symbionts generated from this evolution experiment may correspond to intermediate evolutionary stages towards mutualistic symbionts.

The results obtained in this experiment provide evidences that a strain with a completely different lifestyle (pathogenic and strictly extracellular) can become an intracellular legume symbiont. This conversion relies on the modification (inactivation and recruitment) of biological functions intrinsic to the recipient genome. This process may have been facilitated by the fact that *R. solanacearum* was already adapted to many plant environments and can infect more than 250 different plant species (Genin & Denny, 2012), and thus possibly had a predisposition to symbiosis.

4.3. Biology of the EE protagonists

4.3.1. *Cupriavidus taiwanensis*, the natural symbiont of *M. pudica*

The bacterial species *Cupriavidus taiwanensis* was chosen in the design of the experiment as donor of the symbiotic plasmid containing all the essential symbiotic genes, named pRalta. This rhizobium is a β-proteobacterium and intracellularly infects the indeterminate nodules of the legumes *Mimosa pudica* and *Parapiptadaenia rigida* (Klonowska et al., 2012; Moulin et al., 2001). The strain LMG19424 was isolated in Taiwan from the *Mimosa pudica* legume (Chen et al., 2001b). However, this species is distributed worldwide, being already isolated from *M. pudica* plants in India (Verma et al., 2004) and Costa Rica (Barrett & Parker, 2006), New Caledonia (Klonowska et al., 2012), French Guiana (Mishra et al., 2012) and Chinese mainland (Liu et al., 2011). In *M. pudica* nodules, *C. taiwanensis* bacteroids did not undergo a terminal differentiation and a percentage of 20% of them is able to regrow after the release from the nodules (Marchetti et al., 2011).

The genome *of Cupriavidus taiwanensis* consists in two chromosomes of 3.42 Mb and 2.5 Mb, in addition to a symbiotic plasmid of 0.56 Mb (figure 1.7). Concerning the symbiotic plasmid, *C. taiwanensis* has characteristics of a minimal rhizobium, showing the most compact symbiotic island described so far among rhizobia: a 35 kb fragment containing *nod* and *nif* genes, responsible for nodulation and nitrogen fixation, respectively (Amadou et al., 2008). This symbiotic plasmid harbors two *nod*-boxes. One upstream the *nodBCIJHASUQ* genes possibly arranged in a single operon neighboring the regulatory gene *nodD*, and another one located outside the symbiotic island upstream from the gene *noeM*, which has a role in the

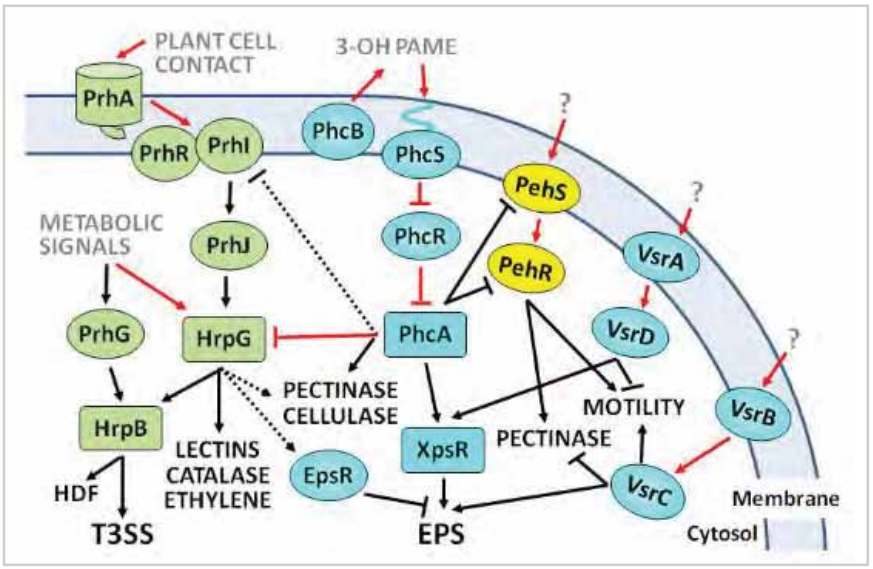


Figure 1.8. Control of virulence functions in *Ralstonia solanacearum* is dependent on a complex regulatory network. The PhcA central regulator controls many of virulence genes. This protein activates EPS and cellulase-encoding genes. On the contrary, it represses swimming motility and T3SS, as well as the transcription of a pectinase-encoding gene via PehR. The transcription of this global regulator is controlled by the *Ralstonia*-specific cell density-dependent quorum sensing system, which demands the production of 3-hydroxy palmitic acid methyl ester or methyl 3-hydroxymyristate (3-OH PAME or 3-OH MAME) by the inner membrane PhcB. At low bacterial cell densities, PhcA expression is repressed, but once the 3-OH PAME/3-OH MAME concentration is increased due to a confinement of bacteria or bacterial cell densities above 10^7 cell/mL, the two-component system PhcS/PhcR is activated. This causes the phophorilation of PhcR, consequently de-repressing the expression of PhcA. Consequently, an improved production of EPS occurs, through the induction of XpsR, which activates the *eps* operon. In the same way, the two-component regulatory systems VsrA/VsrD and VsrB/VsrC are recquired for the activation of *xpsR* transcription. However, VsrD was also shown to affect swimming motility directly by repressing the transcription of flagellum genes. In brief, XpsR integrates VsrAD and PhcA to regulate the *eps* promoter, playing a key role in the control of the biosynthesis of EPS, together with the negative control by EpsR and the positive control by VsrC. Another important system is controled by the regulators HrpB and HrpG, which affect the expression of genes encoding the T3SS and associated effectors. The induction of T3SS genes occurs when a outer membrane receptor, ~~the~~PrhA, perceives a plant signal and transfers it to the membrane-associated proteins PrhI and PrhR. The signal passes then through PrhJ, HrpG and HrpB. In addition, HrpG was shown to control a larger set of T3SS-unrelated genes independent on HrpB, most of them being probably linked to plant-pathogen interactions. Finally, it has been hypothesized that the colonization of the hosts by *R. solanacearum* occurs in a two-step system: firstly, the expression of the T3SS is activated by plant cell contact. At this point, PhcA is not induced due to the low bacterial densities. It allows bacterial motility to be active. Secondly, with the increase in bacterial density in the xylem vessels, the PhcA regulator is induced, activating EPS and cellulase production and repressing T3SS and motility (extracted from Peters et al., 2013).

synthesis of Nod Factors (Daubech et al., 2019). Besides that, *C. taiwanensis* also possess an active T3SS encoded on its chromosome 2. This system does not play any role in the symbiosis with *Mimosa pudica*, but limits the symbiosis with *Leucaena leucocephala* (Saad et al., 2012).

4.3.2. *Ralstonia solanacearum,* a broad host range plant pathogen

The phytopathogen *Ralstonia solanacearum* was chosen as the recipient of the symbiotic plasmid. Like *C. taiwanensis*, it is a β-proteobacterium and is the biological agent behind one of the most important plant bacterial disease worldwide, causing plant wilt in more than 250 plant species, including relevant crops, such as potato, tomato, eggplant and banana. This phytopathogen can infect the hosts via root wounds and sites of emergence of secondary roots. Once inside the plant, the bacteria remains strictly extracellular and can multiply in the cortical cells, then colonize rapidly the xylem vessels, form biofilms and/or spread to the plant aerial parts. General symptoms of the disease include yellowing, wilting and generalized necrosis (Genin and Denny, 2012; Peeters et al., 2013). The genome of this phytopathogen is organized in two circular replicons, one chromosome of 3.7 Mb and one megaplasmid of 2.1 Mb. The chromosome of *R. solanacearum* carries genes classified as essential for basic functions, the so-called housekeeping genes, while the megaplasmid mainly carries genes related to environmental adaptation and virulence, such as secretion systems, flagellar motility and EPS biosynthesis genes (Genin and Denny, 2012).

The pathogenicity of *R. solanacearum* relies on several virulence factors. Among them, a T3SS and a large repertoire of T3 effectors play a major role (Peeters et al., 2013). This system, encoded by the *hrp* genes, is a syringe-shaped apparatus that crosses the cellular membrane to inject Type III effectors (T3Es) in the host cells. In the GMI1000 strain, 70 to 80 T3Es are injected into the plant cytosol in order to favor infection (Erhardt et al., 2010; Tampakaki et al., 2010). Mutants of *R. solanacearum* defective in the T3SS show a non-pathogenic phenotype on its host plants (Peeters et al., 2013). Although the T3Es of *R. solanacearum* are collectively required for virulence, only a subset of *Ralstonia solanacearum* carrying mutations on individual T3Es presented slightly to fully impaired virulence on tobacco, eggplants, tomato and bean depending on the T3E inactivated in each strain (Angot et al., 2006; Lei et al., 2020; Macho et al., 2010). Moreover, T3Es can also induce hypersensitive reactions (HR) on resistant hosts (Asolkar and Ramesh, 2018; Poueymiro et al., 2009; Poueymiro and Genin, 2009).

Additionally, another important factor of virulence of *R. solanacearum* is the biosynthesis of EPS, a polymer of N-acetylgalactosamine that acts in two different ways: this molecule allows the phytopathogen to protect itself against plant defense by hiding bacterial surface features that could be recognized by host plants and during the colonization of the host, the exuded EPS block the plant xylem vessels physically, thus causing the wilting symptoms. The biosynthesis of this molecule depends on the *eps* locus, which is expressed under the negative control of EpsR and the positive control of XpsR and VsrC (Genin and Denny, 2012).

The Type II Secretion System (T2SS) of *R. solanacearum* is also necessary to the virulence of *Ralstonia solanacearum*. This system is responsible for secreting a cocktail of cell-wall-degrading enzymes (such as polygalacturonases, endoglucanase, pectin methylesterase and cellobiohydrolase) that could contribute to the colonization of the plant tissues by the pathogen. Mutants of the genes coding for these enzymes were shown to be less virulent than the wild-type strain and impaired in plant colonization and multiplication within the host (Kang, 1994).

Finally, bacterial motility also seems to be an important determinant of fitness and virulence during the interaction between *R. solanacearum* and its hosts, especially acting in the early stages of host colonization and invasion. Indeed, non-motile and non-flagellated mutants of the GMI1000 strain were shown to have reduced virulence when bacteria are soil-drench inoculated on tomato but not when they are directly inoculated in the plant xylem (Corral et al., 2020). Moreover, hypermotile mutants of this strain were also shown to have reduced virulence (Meng et al., 2011), suggesting that a very precise regulation of bacterial motility is necessary for the full virulence of this plant pathogen. It seems that both swimming and twitching motility are necessary for *R. solanacearum* pathogenicity. However, swimming motility seems to be important for host root colonization and infection, while twitching motility, which is mediated by the bacterial type 4 pilus, is important in both root and xylem colonization (Corral et al., 2020).

The pathogenicity of *Ralstonia solanacearum* is regulated by a complex and interconnected regulatory network (figure 1.8), which includes, among others, the global regulators HrpG, EfpR and PhcA (Genin and Denny, 2012). The regulator HrpG regulates the T3SS of *R. solanacearum* and associated effectors through the activation of the intermediate regulator HrpB, but also 200 other genes independently of HrpB. However, the regulator HrpG is exposed to multiple regulation itself, more specifically by PrhARIJ, which responds to a plant

signals and the PrhG regulator, which responds to a metabolic signal (Marenda et al., 1998; Plener et al., 2010).

PhcA is a central regulator in *R. solanacearum* that controls several hundreds of genes (potentially up to 1500 genes) (Khokhani et al., 2017; Mori et al., 2017; Perrier et al., 2018). Globally, PhcA activates EPS and cellulose-encoding genes and represses swimming motility, T3SS, and siderophore expression, in addition to pectinase-encoding genes via PehR (Genin and Denny, 2012). In addition, this regulator is a metabolic repressor, since PhcA mutants have an increased ability to assimilate nutrient sources and a reduced virulence (Peyraud et al., 2016). The activity of the global regulator PhcA is controlled by a cell-density dependent mechanism specific to *R. solanacearum* and not fully elucidated, which involves the production of 3-hydroxy palmitic methyl ester or methyl 3-hydroxymyristate (3-OH PAME or 3OH MAME) depending on the strains, synthesized by the inner membrane protein PhcB (Genin and Denny, 2012) and the quorum sensing regulatory system PhcSRQ (Schell, 2000; Takemura et al., 2021; Tang et al., 2020). At low cell densities, the activity of PhcA is repressed, while a minimal cell density of 10^7 cells/mL is sufficient to activate PhcA (Schell, 2000).

The virulence regulatory network of *R. solanacearum* also involves the global regulator EfpR. Like PhcA, this regulator is a metabolic repressor in *Ralstonia* and regulates a wide range of genes, such as motility and EPS biosynthesis genes (Perrier et al. 2016). This regulator was shown to act via the activation of a membrane-associated module encoded by the RSc3146-3148 genes (Capela et al., 2017; Perrier et al., 2016). Mutations repressing the gene *efpR* allowed the adaptation of the strain GMI1000 of *R. solanacearum* to bean, which is a distant host of this strain (Guidot et al., 2014).

4.4. Regulatory rewiring of the recipient genome as a main driver of symbiotic adaptation

During the evolution experiment, *Ralstonia* clones rapidly acquired (Marchetti et al., 2010) and then strongly improved their symbiotic capacities (Marchetti et al., 2017; Marchetti et al., 2014). In 16 evolution cycles, the team observed a strong improvement in nodulation and intracellular infection of nodules. In order to understand the molecular mechanisms behind the phenotypic evolution, the genome of several evolved clones was resequenced and mutations were correlated to the symbiotic phenotype of these clones. This approach allowed the identification of adaptive mutations responsible for the main phenotypic shifts (Capela et al., 2017; Guan et al., 2013; Marchetti et al., 2010; Tang et al., 2020) (figure 1.9). The first

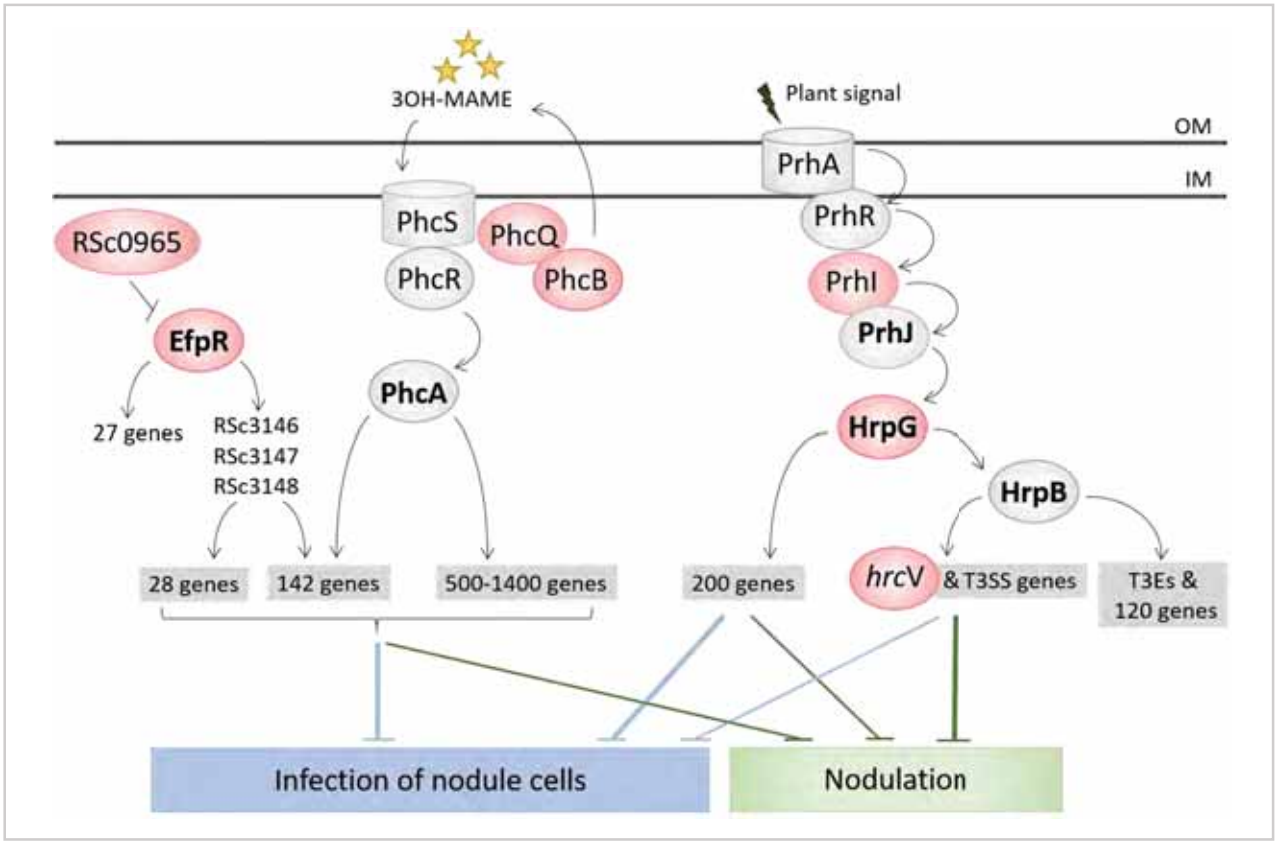


Figure 1.9. Adaptation to endosymbiosis mainly occurred through regulatory rewiring in the evolution experiment. Genes mutated in the experiment are indicated in red. Nodulation was gained in one nodulating ancestor by inactivation of the T3SS via a stop mutation in *hrcV* in contact with the plant host. Two nodulating ancestors were obtained, as well as a first level of intracellular infection, via the inactivation of the virulence regulator HrpG. Following the first level of nodule intracellular infection, the optimization of this trait was reached either by inactivation of the EfpR pathway (in the lineages B and G, with the mutations up^{115}-RSc0965 and *efpR*-E66K, respectively) or a modulation of the quorum sensing system controlling PhcA. EfpR and PhcA controls a common set of genes as well as other specific genes. Transcriptional regulators are in bold. 3OH-MAME, (*R*)-3-hydroxymyristic acid methyl ester, quorum sensing molecules produced by the GMI1000 strain of *R. solanacearum* (extracted from Doin de Moura et al., 2020).

acquired symbiotic trait, the ability to enter legume roots and form nodules was the result of the inactivation of the T3SS, the main virulence factor of *Ralstonia solanacearum*. A stop mutation in *hrcV* encoding a structural gene of the T3SS occurred in one nodulating ancestor, while stop mutations in the regulator *hrpG* occurred in the two other nodulating ancestors. However, these three ancestors were not symbiotically equivalent. Indeed, nodules formed by the *hrcV* mutant were necrotic and only extracellularly infected whereas nodules formed by *hrpG* mutants were partially intracellularly invaded (Marchetti et al., 2010). The ability to intracellularly infect nodule cells was subsequently acquired in some lineages derived from the *hrcV*-mutated ancestor either through a frameshift mutation in *vsrA* encoding another virulence regulator, or mutations in the promoter region of *hrpG* or mutations in *prhI*, encoding the sigma factor acting upstream from HrpG in the regulatory cascade (Guan et al., 2013). However, nodules formed by these evolved clones had only a small number of infected cells and necrotic zones. Later in the evolution cycles, the improvement in the mean number of infected cells and the absence of necrosis in the nodules were obtained through mutations affecting directly or indirectly the global regulators EfpR or PhcA. In five different evolved lineages, either mutations in the gene *efpR* itself or the intergenic mutation upstream from a gene of unknown function triggered a constitutive repression of the transcriptional regulator EfpR (Capela et al., 2017). In two other lineages, a modulation of the quorum threshold activating PhcA generated by mutations in two components of the Phc quorum sensing system, *phcB* and *phcQ,* also allowed the optimization of the intracellular infection of nodules (Tang et al., 2020). Both EfpR and PhcA play central roles in the *R. solanacearum* virulence and metabolism (Capela et al., 2017; Perrier et al., 2018; Perrier et al., 2016; Peyraud et al., 2016). Moreover, these two regulons are overlapping. Whether common and/or specific EfpR- and PchA-regulated genes are involved in intracellular infection improvement remains to be determined.

Altogether, data accumulated in the team showed that the rewiring of regulatory circuits seems to be the main driver of adaptation of the phytopathogen *Ralstonia solanacearum* to symbiosis with the legume *Mimosa pudica* (figure 1.9). Indeed, all identified adaptive mutations identified so far, except *hrcV*, affect regulatory genes. This is in line with other evolution experiments showing that global regulators are primary targets of evolution (Carroll et al., 2015; Hindré et al., 2012; Pankey et al., 2017; Philippe et al., 2007). This is probably

because, changing the expression of several genes in a concerted way is required to achieve important evolutionary leaps.

4.5. Discovery of a hypermutagenesis mechanism that accelerates HGT-based evolution

Bacteria adapt to new environments and conditions through the generation of new mutations. Mutations in genes involved in DNA repair can generate hypermutator strains, characterized by a constitutive elevation of mutation rates. Interestingly, hypermutator genotypes were already shown to be selected in evolution experiments or during host colonization (Denamur and Matic, 2006; Raynes and Sniegowski, 2014). Besides the constitutive hypermutators, transient mutators of bacteria can be generated in response to stress *via* a mechanism known as stress-induced mutagenesis (SIM) (Foster, 2007; Galhardo et al., 2007). In this process, the expression of error-prone DNA polymerases under the control of major stress-induced regulatory pathways such as the SOS response or RpoS is responsible for the hypermutator phenotypes of the resulting strains. Interestingly, the sequencing of evolved clones generated during this evolution experiment revealed a large number of mutations arising in each evolution cycle, even if the clones were not constitutive hypermutators. Subsequent investigations indicated that bacterial clones undergo a transient hypermutagenesis phase (~5–10× increase in mutation rate) during each exposure to the plant culture medium, with a further increase (~4×) in the presence of *Mimosa pudica* plantlets. This phenomenon is due to the presence of error-prone DNA polymerases from the *imu*ABC family encoded by the symbiotic plasmid pRalta. In the same work, by re-evolving clones carrying the *imu*ABC genes and their isogenic *imu*ABC mutants for 5 cycles, the team observed that SIM accelerates adaptation in this evolution experiment (Remigi et al., 2014). Interestingly, more than half of rhizobial symbiotic plasmids carry error-prone DNA polymerases. Due to this observation, the team has hypothesized that the co-transfer of the essential symbiotic genes with the *imuABC* mutagenic cassette may facilitate the adaptation of bacteria to the new symbiotic lifestyle. By promoting genetic diversification of the rhizopsheric bacterial population, potential symbiotic efficient variants can appear and be selected by the plant. However, on the long term, the phenomenon of hypermutagenesis may be detrimental for the bacterial fitness, since some accumulated mutations may have deleterious effects (Denamur and Matic, 2006). This may lead, once bacteria are well adapted, to a selection against hypermutagenesis, finally modulating the mutation rates, a phenomenon that was already described for constitutive

hypermutators in several evolution experiments (McDonald et al., 2012; Swings et al., 2017; Wielgoss et al., 2013).

5. PhD project

Before this thesis, the analysis of this EE was centered on the molecular and phenotypic characterization of individual evolved clones, which gave only a partial and non-dynamic view of the adaptive process. Moreover, the main adaptive mutations identified so far affected global regulators controlling positively and negatively hundreds of genes, hindering the identification of the functions involved in symbiotic adaptation. The main objective of this thesis was to analyze this experiment in a more dynamic and exhaustive way to better understand the genetic bases of adaptation to legume symbiosis and the selective forces having driven the emergence of the new symbionts. For this purpose, we used a metagenomic approach coupled to extensive phenotype/genotype correlation analyses to describe evolution at the molecular, phenotypic and genetic levels. This first part of the work allowed the detection of mutations that were rapidly fixed in evolved populations. Among them, a pool of new adaptive mutations were identified and their associated symbiotic phenotypes determined, revealing the symbiotic trait predominantly selected in this experiment. In a second part of the work, we functionally characterized the most adaptive of these mutations through a battery of phenotypic assays, trying to get insights into the biological functions associated with symbiotic adaptation. Finally, in a third part, we analyzed the EE conducted up to cycle 60. Although mutualism was not yet achieved at that stage, we evidenced the mutational convergence in one gene, RSc2277, allowing the detection of low levels of nitrogenase activity. This work opens the way for a better understanding of the selective forces acting on the emergence of legume symbionts as well as the biological functions involved in, or interfering with, the establishment of symbiosis while pointing out the difficulty to evolve mutualistic nitrogen-fixing bacteria belonging to a non-rhizobium genus

CHAPTER 2

A selective bottleneck during host entry drives the evolution of new legume symbionts

This chapter was submitted to the journal PLoS Genetics and is currently under review

This manuscript was also deposited in the pre-print server bioRxiv

https://doi.org/10.1101/2022.03.03.482760

A selective bottleneck during host entry drives the evolution of new legume symbionts

Ginaini Grazielli Doin de Moura[1], Saida Mouffok[1], Nil Gaudu[1], Anne-Claire Cazalé[1], Marine Milhes[2], Tabatha Bulach[2], Sophie Valière[2], David Roche[3], Jean-Baptiste Ferdy[4], Catherine Masson-Boivin[1], Delphine Capela[1] and Philippe Remigi[1]

[1]LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France

[2]INRAE, US1426, GeT-PlaGe, Genotoul, Castanet-Tolosan, France

[3]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France

[4]CNRS, UMR5174 EDB, Université Toulouse 3 Paul Sabatier, Toulouse, France.

Email for correspondence: delphine.capela@inrae.fr, philippe.remigi@inrae.fr

Abstract

During the emergence of new host-microbe symbioses, multiple selective pressures –acting at the different steps of the microbial life cycle– shape the phenotypic traits that jointly determine microbial fitness. However, the relative contribution of these different selective pressures to the adaptive trajectories of microbial symbionts are still poorly known. Here we characterized the dynamics of phenotypic adaptation and its underlying genetic bases during the experimental evolution of a plant pathogenic bacterium into a legume symbiont. We observed that fast adaptation was predominantly driven by selection acting on competitiveness for host entry, which outweighed selection acting on within-host proliferation. Whole-population sequencing of evolved bacteria revealed that phenotypic adaptation was supported by the continuous accumulation of new mutations (fuelled by a transient hypermutagenesis phase occurring at each cycle before host entry, a phenomenon described in previous work) and the sequential sweeps of cohorts of mutations with similar temporal trajectories. The identification of adaptive mutations within the fixed mutational cohorts showed that several adaptive mutations can co-occur in the same cohort. Moreover, all adaptive mutations improved competitiveness for host entry, while only a subset of those also improved within host proliferation. Computer simulations predict that this effect emerges from the presence of a strong selective bottleneck at host entry occurring before within host proliferation and just after a hypermutagenesis phase in the rhizosphere. Together, these results show how selective bottlenecks can alter the relative influence of selective pressures acting during bacterial adaptation to multistep infection processes.

Introduction

Many bacterial lineages have evolved the capacity to establish symbiotic associations, either beneficial, neutral or parasitic, with eukaryotic hosts. These interactions form a dynamic continuum along which bacteria can move (Drew et al., 2021). Lifestyle changes may arise due to ecological (resource availability, host environment changes or host shifts) or genomic (mutations, acquisition of new genetic material) modifications. These new interactions are often initially sub-optimal for the bacterial partner (Bonneaud and Longdon, 2020; Longdon et al., 2015; Nandasena et al., 2007), which may then adapt to the selective pressures associated with its new life-cycle. For instance, horizontally-transmitted microbes alternate between at least two different habitats, the host and the environment, where they will face a

variety of constraints and selective pressures (Obeng et al., 2021; Robinson et al., 2019). Biphasic life cycles therefore require bacteria not only to be able to enter and exit their hosts, but also to replicate and persist within each habitat, which entails adapting to abiotic stressors, host immunity, competitors, as well as being able to use specific nutrient sources. However, the relative influence of the different selective pressures on the dynamics and the trajectory of bacterial adaptation to a new interaction is poorly known.

Rhizobia are examples of facultative host-associated bacteria that can either live freely in soil or in symbiotic mutualistic associations with legume plants (Wheatley et al., 2020). In most legumes, rhizobia penetrate the root tissue through the formation of so-called infection threads (ITs). Most of the time, only one bacterium attaches to the root hair and initiates the formation of Its (Gage, 2002), thus creating a strong population bottleneck between the rhizosphere and internal root tissues. Bacteria then divide clonally within ITs, while at the same time a nodule starts to develop at the basis of the infected root hair. Hundreds of millions of bacteria are then released from ITs inside the cells of the developing nodule and differentiate into nitrogen fixing bacteroids (Coba de la Peña et al., 2017; Gage, 2004). After several months, in nature, nodule senescence leads to the release of a part of bacterial nodule population in the surrounding soil. To fulfil their life cycle, rhizobia rely on the activity of numerous bacterial genes, allowing signal exchanges with the host plant, proliferation and metabolic exchanges within nodule cells, while maintaining free-living proficiency (Wheatley et al., 2020). Somewhat surprisingly given their complex and specialized lifestyle, rhizobia evolved several times independently. Horizontal transfer of key symbiotic genes is a necessary, though often not sufficient, condition for a new rhizobium to emerge (Abe et al., 1998; Hirsch et al., 1984a; Marchetti et al., 2010; Nakatsukasa et al., 2008). Additional steps of adaptation of the recipient bacterium, occurring during evolution under plant selection, are believed to be needed to actualize the symbiotic potential of emerging rhizobia when the recipient genotype is not already compatible with legume symbiosis (Masson-Boivin et al., 2009).

In a previous study, we used experimental evolution to convert the plant pathogen *Ralstonia solanacearum* into an intracellular legume symbiont (Marchetti et al., 2017). In this experiment, we transferred a symbiotic plasmid (pRalta) from the rhizobium *Cupriavidus taiwanensis* LMG19424, the *Mimosa pudica* symbiont, into the GMI1000 strain of *Ralstonia solanacearum*. We first obtained three nodulating variants, two infecting nodules

intracellularly (CBM212, CBM349) and one infecting nodules extracellularly (CBM356) (Marchetti et al., 2010), which were used as ancestors to evolve 18 parallel bacterial lineages through 16 serial cycles of nodulation on *Mimosa pudica* plants (Marchetti et al., 2017). Some highly adaptive mutations involved in the acquisition of nodulation and intracellular infection of nodules were previously identified (Capela et al., 2017; Guan et al., 2013; Marchetti et al., 2010; Tang et al., 2020). However, these previous works only gave us a partial and non-dynamic view of the adaptive process at both phenotypic and molecular levels. Here, we analysed the dynamics of phenotypic and molecular evolution in 5 lineages evolved for 35 cycles of nodulation. We used whole-population genome sequencing and examined the selective and genetic bases of adaptation. Adaptation proceeded rapidly during the first cycles of evolution, and was underpinned by the fixation of successive cohorts of mutations within populations. We then identified adaptive mutations in two lineages and evaluated their effect on each symbiotic stage of the rhizobial life cycle. Our experimental data indicated that selection for nodulation competitiveness outweighs selection for multiplication within host. Computer simulations further showed that the selective bottleneck at host entry and the chronology of symbiotic events are critical drivers of this evolutionary pattern.

Results

Fast adaptation of new legume symbionts during the first cycles of evolution

To investigate the genetic and evolutionary conditions that promote the evolution of new rhizobia, we previously experimentally evolved the plant pathogen *R. solanacearum* GMI1000 into intracellular *M. pudica* symbionts through 16 serial cycles of nodulation on *M. pudica* plants (Marchetti et al., 2017). Here we continued the evolution of five lineages (referred to as lineages B, F, G, K, and M) until cycle 35 (Fig 2.1a, S2.1 Fig and S2.1 Table). After 35 cycles, no nitrogen fixation sustaining plant growth was observed (S2.2 Fig). Therefore, in this manuscript, we focussed exclusively on the analysis of changes in bacterial symbiotic fitness (estimated for each strain as the frequency of bacteria present in nodule pools collected from 10 plants relative to a competing reference strain and normalized by the frequency of each strain in the inoculum). To analyse the dynamics of fitness changes over time, we compared the relative fitness of the nodulating ancestors and evolved clones from cycles 16 and 35 by replaying one nodulation cycle in competition with the natural symbiont *C. taiwanensis.* Fitness trajectories were similar in the five lineages. While the three nodulating ancestors

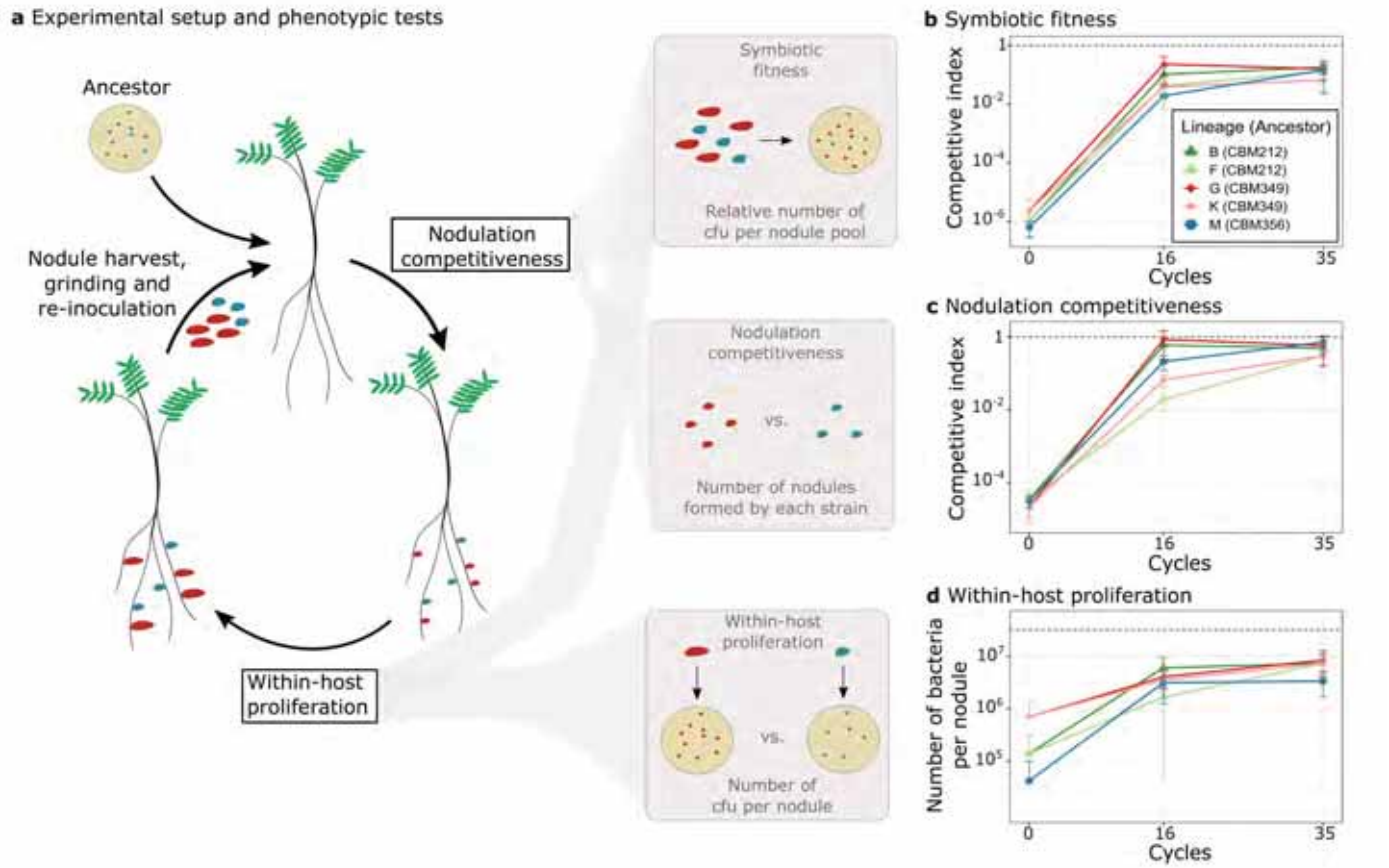


Fig 2.1. Evolution of the symbiotic properties of *Ralstonia* clones along evolution cycles.
a, Left: overview of the evolution cycles showing the symbiotic steps determining bacterial fitness (nodulation competitiveness and within-host proliferation) and human intervention (nodule harvest, grinding and re-inoculation). Nodule colors reflect bacterial genotypes: blue represents the ancestor, and red represents a mutant with increased fitness. Right: schematic representation of the main phenotypic measurements performed in this work, where the symbiotic phenotypes of one evolved strain can be compared to those of a reference strain *(e.g*. the reference symbiont *C. taiwanensis*). Competitive indexes for symbiotic fitness are calculated as the ratio of evolved *vs*. reference clones in nodule bacterial populations normalized by the ratio of strains in the inoculum. Competitive indexes for nodulation competitiveness are calculated as the ratio of nodules formed by each strain normalized by the inoculum ratio. Within-host proliferation is measured in independent single inoculations of each strain. **b-d**, Relative symbiotic fitness (**b**), nodulation competitiveness (**c**) and within-host proliferation (**d**) of nodulating ancestors (cycle 0) and evolved clones isolated from cycles 16 and 35 were compared to *C. taiwanensis*. Values correspond to means ± standard deviations. Data were obtained from at least three independent experiments. For each experiment, nodules were harvested from 10 plants (**b**), 20 plants (**c**) and 6 plants (**d**). The sample size (n) is equal to n=3 (**b**,**c**) or comprised between n=15-18 (**d**). Raw data are available in Supplementary Table 2. cfu: colony-forming units. The data underlying panels b–d can be found in S1 Table.

were on average 10^6 times less fit than *C. taiwanensis*, fitness improved very quickly during the first 16 cycles, then slower until cycle 35 (Fig 2.1b). On average, evolved clones isolated at cycle 16 and cycle 35 were 47 and 19 times less fit than *C. taiwanensis*, respectively (Fig 2.1b). Because bacterial symbiotic fitness in our system depends on both the capacity of strains to enter the host and induce nodule formation (nodulation competitiveness) and to multiply within these nodules (within-host proliferation), we analysed how each of these two fitness components changed during the experiment (Fig 2.1c-d). In the five lineages, evolutionary trajectories of nodulation competitiveness resembled that of fitness with fast improvement during the first 16 cycles, and a slow-down during the last cycles. On average, nodulating ancestors were $5x10^4$ times less competitive than *C. taiwanensis*, while cycle 16 and cycle 35 evolved clones were only 17 and 4 times less competitive than *C. taiwanensis*, respectively (Fig 2.1c). Moreover, evolved clones B16, B35, G16, G35 and M35 were not statistically different from *C. taiwanensis* in terms of nodulation competitiveness. Within-host proliferation also improved mostly during the first 16 cycles in all lineages. The rate of improvement then slowed down in two lineages and continued to increase significantly in the three others. As previously published, the three nodulating ancestors display different capacities to infect nodules (Marchetti et al., 2010; Marchetti et al., 2017), the extracellularly infective ancestor CBM356 being the less infective (750 times less than *C.taiwanensis*) and the intracellularly infective CBM349 being the most infective (45 times less than *C. taiwanensis*). On average, cycle 16 and cycle 35 evolved clones were 8 and 5 times less infective than *C. taiwanensis* (Fig 2.1d).

Altogether, the symbiotic properties of evolved clones improved rapidly during the first cycles of the evolution experiment and then adaptation slowed down. Both symbiotic traits, nodulation competitiveness (*i.e*. host entry) and within-host proliferation, improved greatly. However, gains in nodulation competitiveness (average factor of 11,500 between ancestors and cycle 35 clones) were *ca.* 150 times higher than gains in proliferation (average factor of 74 between ancestors and cycle 35 clones). Moreover, although the difference in nodulation competitiveness between *C. taiwanensis* and the nodulating ancestors was much more important ($>10^4$ fold) than the difference in proliferation ($<10^3$ fold), nodulation competitiveness of evolved clones reached the level of *C. taiwanensis* in three lineages (B, G and M) while within-host proliferation remained lower than *C. taiwanensis* in all of them after 35 cycles.

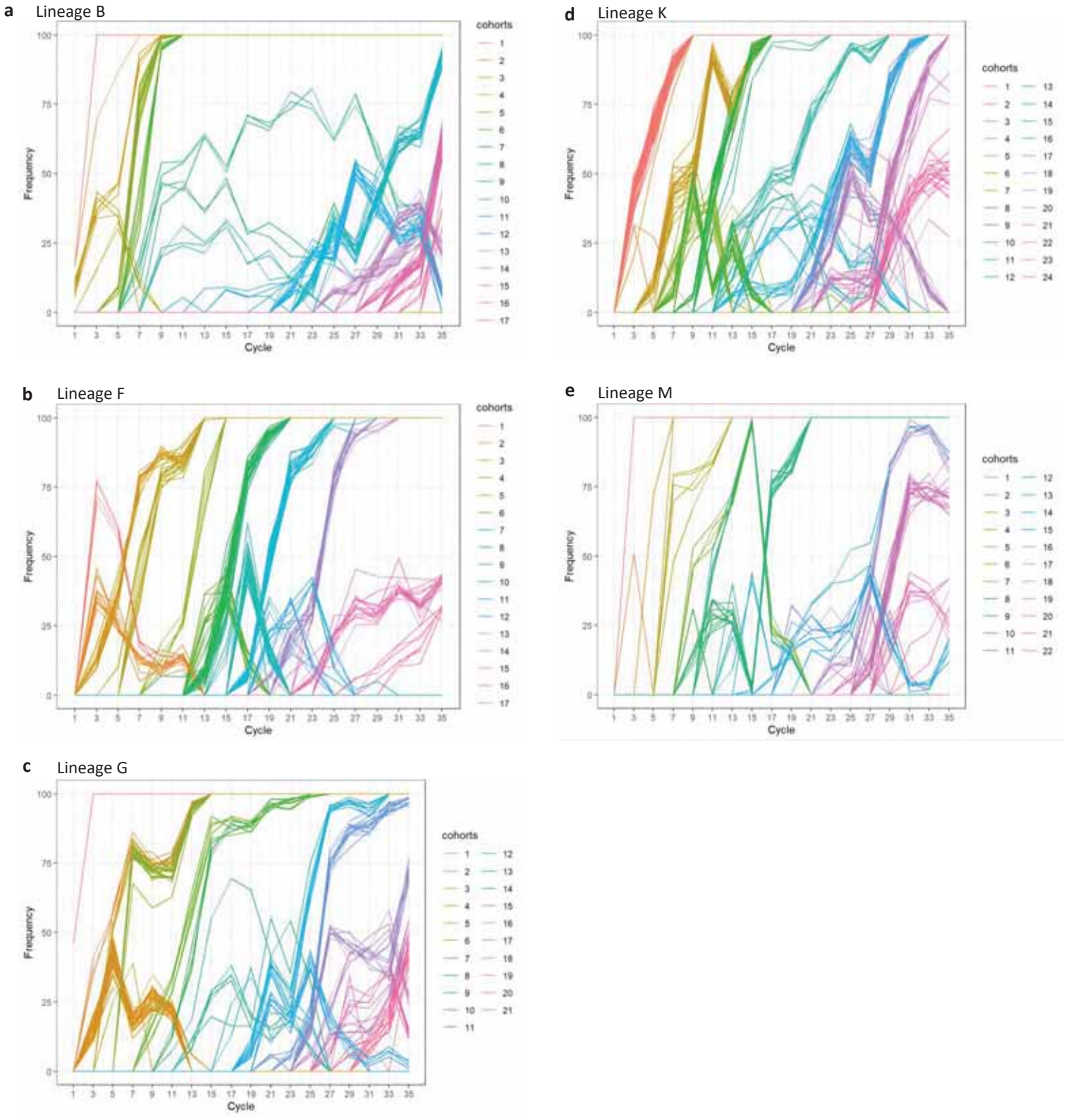


Fig 2.2. Dynamics of molecular evolution. Allele frequency trajectories of mutations that attained a frequency of 30% in at least one population of the B, F, G, K and M lineages (**a**,**b**,**c**,**d**,**e**). Mutations with similar trajectories were clustered in cohorts and represented by different colors. For simplicity, mutations travelling alone were also called cohorts. The data underlying panels a–e can be found in S4 Table.

The dynamics of molecular evolution is characterized by multiple selective sweeps of large mutational cohorts

The rapid adaptation in this experiment was unexpected given that strong population bottlenecks, such as the one occurring at the nodulation step, limit effective population size at each cycle (see S2.1 Table for inoculum population size and effective population size) and are generally known to limit the rate of adaptive evolution (Moxon and Kussell, 2017). To describe the dynamics of evolution at the molecular level, we performed whole-population sequencing of evolved lineages using the Illumina sequencing technology. Populations of the five lineages were sequenced every other cycle until cycle 35 with a minimum sequencing coverage of 100x and a median of 347x. We detected a very large number of mutations in all sequenced populations. In the five lineages, a total of 4114 mutations were detected above a frequency of 5%, a threshold below which we considered that the detected mutations are potentially errors. This corresponds to 382 to 1204 mutations per lineage with 23.6 new mutations per cycle on average (S2.3-S2.4 Tables). This high number of mutations is a result of a previously characterized transient hypermutagenesis phenomenon occurring in the rhizosphere of *M. pudica* (Remigi et al., 2014). This phenomenon is due to the presence of a cassette of error-prone DNA polymerases on the symbiotic plasmid pRalta, whose expression is induced under stressful conditions and is regulated by the SOS response. Hypermutation was specifically observed in the rhizosphere of *M. pudica* but not in nodules nor in bacterial cultures in rich medium. In evolved populations, mutations accumulated throughout the experiment, showing no sign of slowing down until the end of the experiment, which suggests that no anti-mutator mutations established in these populations as we might have expected (Lynch et al., 2016). We evaluated genetic parallelism among mutations detected above a frequency of 5% in the populations by computing G scores (Tenaillon et al., 2016), a statistics used to point out genes that could be mutated more often than expected by chance. In spite of high mutation rates that may obscure signals of genetic convergence, we detected signatures of parallelism at the gene level among our list of mutations (observed sum of G scores of 7,540.5, compared to a mean sum of 5,401.01 after 1,000 randomized simulations, $Z = 31.23$, $P < 10^{-200}$). The randomized simulations identified 171 genes with a higher number of mutations than expected by chance (Bonferroni adjusted *P* value<0.01).

To simplify the analysis of mutational trajectories, we then focused on the 819 mutations (out of the total of 4414 detected mutations) that rose above a frequency of 30%. These mutations

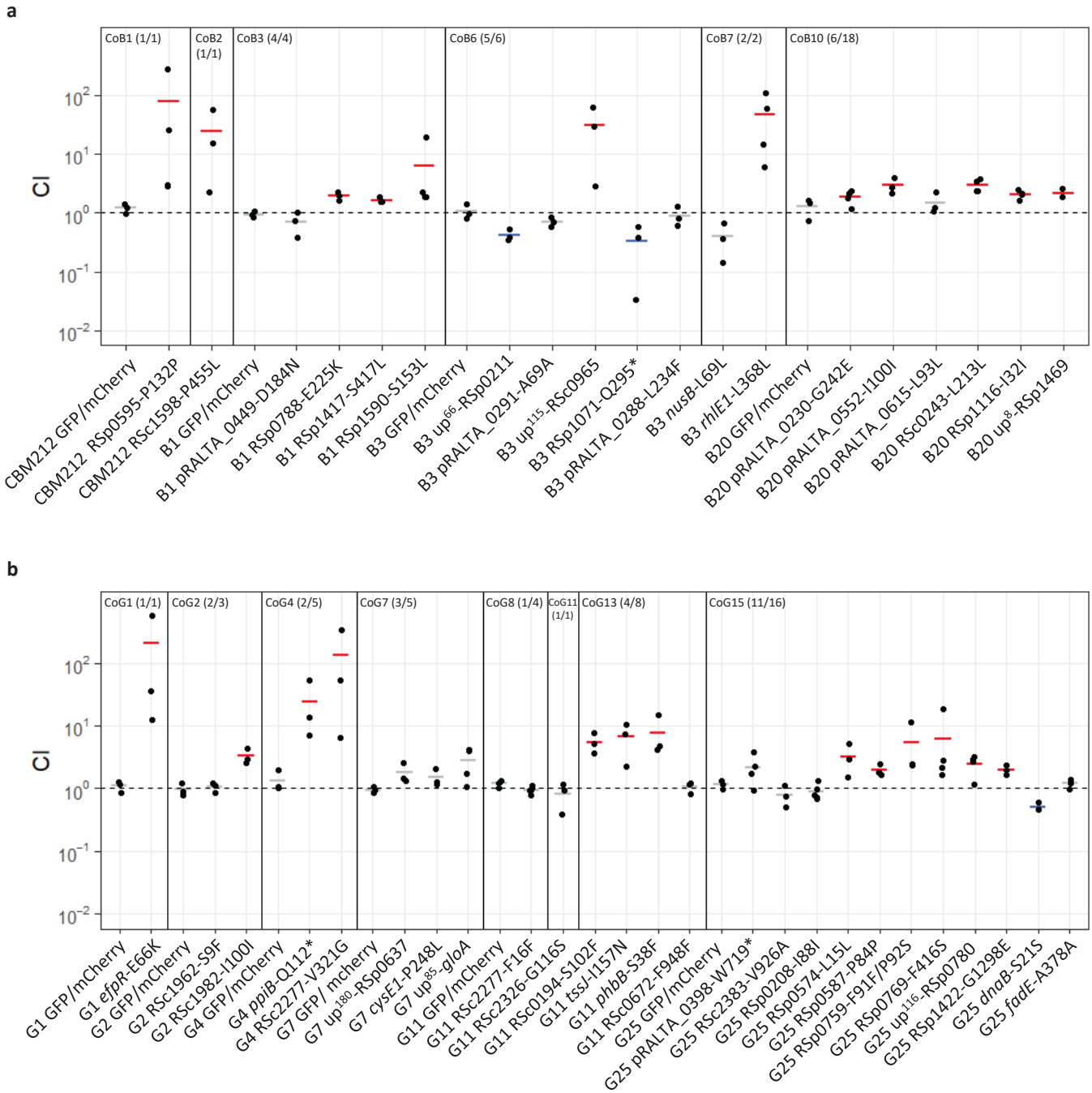


Fig 2.3. Symbiotic fitness of reconstructed mutants from lineages B and G.
Relative symbiotic fitness of evolved clones carrying reconstructed mutations from fixed mutational cohorts identified in lineages B (**a**) and G (**b**). Evolved clones in which mutations were reconstructed are indicated at the bottom of the graph. Vertical lines separate the different mutational cohorts. Cohort number and the number of tested mutations from the cohort on the total number of mutations present in the cohort are indicated in brackets. Competitive indexes (CI) were calculated as the ratio of the mutant strain on the isogenic parental strain in bacterial nodule populations normalized by the inoculum ratio. GFP/mCherry correspond to control co-inoculation experiments of strains derived from the same evolved clone labelled with different fluorophore either GFP or mCherry. Horizontal segments correspond to mean values of CI. Red segments indicate significantly beneficial mutations, gray segments indicate neutral mutations and blue segments indicate significantly deleterious mutations ($P<0.05$, *t*-test with Benjamini-Hochberg correction). Data were obtained from 3 to 5 independent experiments. The sample size (n) is comprised between n=3-5. Each CI value was obtained from pools of 40 to 227 nodules harvested from 10 plants. The data underlying panels a–b can be found in Doin de Moura et al., 2022 (Figure 3 - source data I, doi: https://doi.org/10.1101/2022.03.03.482760).

did not show independent trajectories. Instead, we observed that groups of mutations arose synchronously and showed correlated temporal trajectories, enabling us to cluster them in cohorts (Buskirk et al., 2017; Lang et al., 2013) (Fig 2.2 and S2.3 Table). Cohorts increased in frequency with variable speed, some being fixed within 1 to 3 cycles (representing *ca*. 25-75 bacterial generations) while others reached fixation in up to 21 cycles (representing more than 500 bacterial generations). Fixed cohorts can be very large (up to 30 mutations) (S2.3 Table), which suggests hitchhiking of neutral or slightly deleterious mutations with one or several beneficial mutations acting as driver (or co-drivers) of the cohort (Lang et al., 2013; Nguyen Ba et al., 2019). Among the mutations that rose above 30% frequency at some stage of the experiment, 35% later declined until extinction. This is indicative of clonal interference, *i.e.* the co-occurrence of multiple subpopulations competing against each other within populations (McDonald, 2019), which can readily be observed on Muller plots for lineages B and G (S2.3 Fig).

Overall, in these 5 lineages, the pattern of molecular evolution was characterized by a steady accumulation of mutations along the successive cycles and the formation of mutational cohorts, some of which containing a large number of mutations. Despite strong population bottlenecks at the root entry, mutation supply was not limiting as evidenced by the co-occurrence of competing adaptive mutations. Yet, multiple (and sometimes rapid) selective sweeps occurred throughout the 35 cycles in all lineages, suggesting that strong selection is acting on these populations.

Large fixed cohorts contain multiple adaptive mutations

In our system, strong nodulation bottlenecks lead to a small effective population size and to the possibility that genetic drift might be responsible for the observed allelic sweeps. To determine whether drift or adaptation are causing these sweeps, we searched for adaptive mutations in fixed (or nearly fixed, >90% frequency) cohorts from the two lineages that have the highest symbiotic fitness after 35 cycles: B and G. To do so, we introduced (by genetic engineering) individual mutations from each cohort of interest into an evolved clone carrying all (or, when not available, almost all) previously fixed cohorts. This allowed us to test the symbiotic fitness effect associated with each mutation in a relevant genetic background, taking into account possible epistatic effects arising from mutations that were previously acquired in this clone. Competitive index (ratio of strains carrying the mutant vs. the wild-type

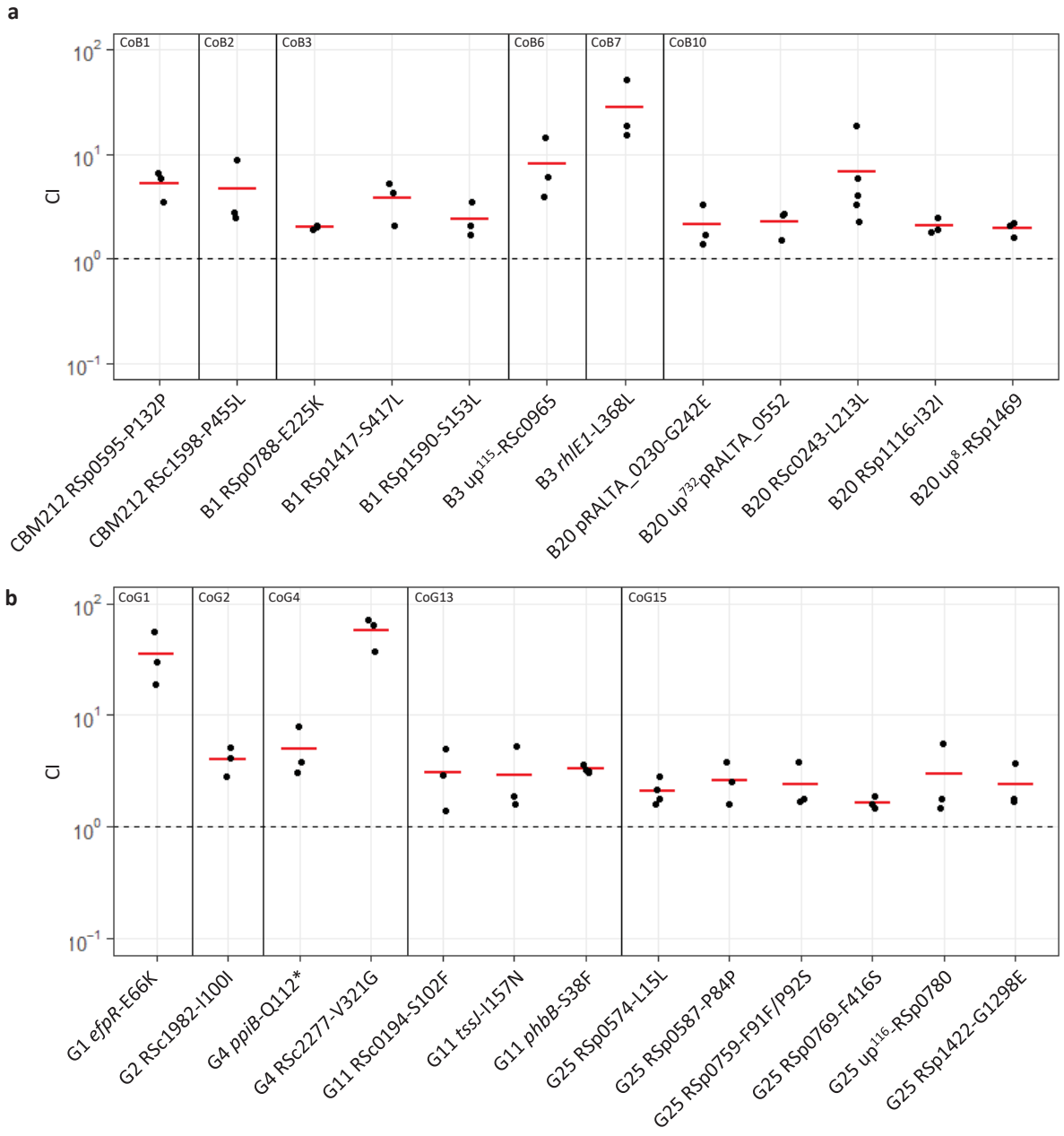


Fig 2.4. Nodulation competitiveness of adaptive mutants

Nodulation competitiveness effect of adaptive mutants from lineages B (**a**) and G (**b**). Evolved clones in which mutations were reconstructed are indicated at the bottom of the graph. Vertical lines separate the different mutational cohorts. Cohort number and the number of tested mutations from the cohort on the total number of mutations present in the cohort are indicated in brackets. CI were calculated as the ratio of the number of nodules formed by the mutant strain on the number of nodules formed by the isogenic parental strain normalized by the inoculum ratio. Horizontal segments correspond to mean values of CI. Red segments indicate significantly beneficial mutations (P<0.05, t-test with Benjamini-Hochberg correction). Data were obtained from 3 to 4 independent experiments. The sample size (n) is comprised between n=3-4. Each CI value was obtained from 95-98 nodules harvested from 20 plants. The data underlying panels a–b can be found in Doin de Moura et al., 2022 (Figure 4 - source data I, doi: https://doi.org/10.1101/2022.03.03.482760).

alleles in nodule pools harvested from 10 plants normalized by the inoculum ratio) were measured for 44 mutations belonging to 14 cohorts from the two lineages (Fig 2.3a-b). Twenty-five mutations significantly increased the fitness of evolved clones and were thus beneficial for symbiosis while 14 mutations were neutral and 4 were slightly deleterious for symbiosis (Fig 2.3a-b, Table 2.1 and S2.4 Table). Among the 44 genes targeted by the reconstructed mutations, 8 have a higher number of mutations than expected by chance and adaptive mutations were identified in 5 of them (Table 2.1).

Several interesting results emerge from this data set. First, some mutations show very strong adaptive effects, improving more than 100 times the fitness of bacteria, in particular during the first cycles of evolution. Second, many cohorts carry more than 1 adaptive mutation: five cohorts contain between 2 and 6 adaptive mutations. Third, 36% of adaptive mutations were found to be synonymous mutations, confirming that these mutations can play a significant role in adaptive evolution (Agashe, 2021; Bailey et al., 2021). All adaptive synonymous mutations except one converted frequently used codons in *R. solanacearum* into unusual ones (S2.5 Table), which may explain their functional effects. Moreover, when inspecting gene annotations, we found that the 25 beneficial mutations target various biological functions (Table 1). Strikingly, nearly 40% affect regulatory functions such as global transcription regulators (*efpR* and Rsc0965) (Capela et al., 2017), putative quorum quenching (RSp0595), signal transduction (RSc1598), unknown transcription regulator (RSc0243), protein dephosphorylation (Rsp1469), protein folding (*ppiB*), degradosome (*rhlE1*), and DNA methylation (RSc1982). Another significant proportion of mutations (28%) affect metabolism and transport genes (*phbB*, RSc0194, RSc2277, Rsp1116, RSp1417, RSp1422 and RSp1590) and 3 mutations are located in genes belonging to the type VI secretion system (T6SS) operon (*tssJ*, RSp0759, Rsp0769). The observed genetic architecture (*i.e*. the large number of genes and molecular functions targeted by adaptive mutations) indicates that symbiotic traits are complex and constitute a large mutational target for adaptive evolution.

Adaptive mutations predominantly improve nodulation competitiveness

Next, we measured the effect of each adaptive mutation on the two traits contributing to the total symbiotic fitness, the nodulation competitiveness (*i.e.* the capacity of bacteria to enter the root and induce nodule formation) and within-host proliferation (*i.e.* the capacity of bacteria to multiply within each nodule), in order to uncover which of these two symbiotic

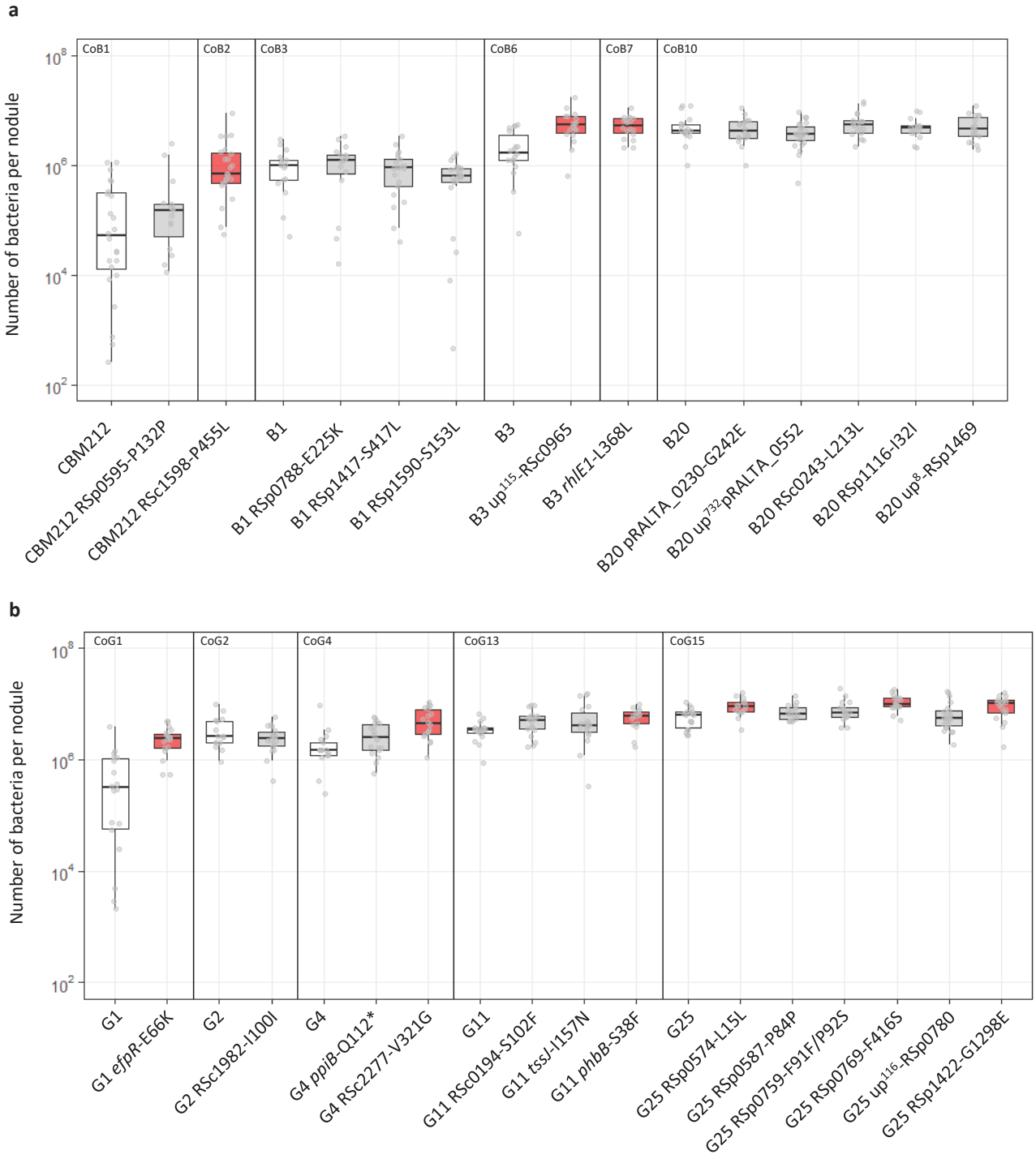


Fig 2.5. Within-host proliferation of adaptive mutants and corresponding isogenic parental clones

Distribution of the number of bacteria per nodule recovered for adaptive mutants from lineage B (**a**) and G (**b**) and their corresponding isogenic parental clones at 21 dpi. Rectangles span the first quartiles to the third quartiles, bold segments inside the rectangle show the median, unfiled circles represent outliers, whiskers above and below each box show the largest and smallest values within 1.5 times interquartile range above and below the 75th and 25th percentile, respectively. Red boxes indicate the mutations that significantly improved infectivity compared to the parental evolved clone (P<0.05, Wilcoxon test with the Benjamini-Hochberg correction) shown in the white boxes. Data were obtained from 3 to 5 independent experiments. For each experiment, nodules were harvested from 6 plants. The sample size (n) is comprised between n=15-24. The data underlying panels a–b can be found in Doin de Moura et al., 2022 (Figure 5 - source data I, doi: https://doi.org/10.1101/2022.03.03.482760).

traits can explain the observed fitness gain. All mutations increasing symbiotic fitness improved nodulation competitiveness of evolved clones (Fig 2.4a-b). Gains in nodulation competitiveness ranged from 1.6 to 58-fold, with three mutations providing the highest gains: RSc2277-V321G (58-fold), *efpR*-E66K (35-fold) and *rhlE1*-L368L (28-fold). By contrast, within-host proliferation was improved by only a subset of adaptive mutations (Fig 2.5a-b). Only 3 mutations from lineage B (RSc1598-P455L, up^{115}-RSc0965 and *rhlE1*-L368L) and 6 mutations from lineage G (*efpR*-E66K, RSc2277-V321G, *phbB*-S38F, RSp0574–L15L, RSp0769-F416S and RSp1422-G1298E) significantly improved this fitness component. Moreover, gains in *in planta* proliferation were generally lower than gains in nodulation competitiveness and ranged from 1.6 to 6.4-fold. The highest proliferation gains were produced by the five mutations RSc1598-P455L (6.4-fold), *efpR*-E66K (3.7-fold), up^{115}-RSc0965 (2.6-fold), RSc2277-V321G (2.5 fold) and *rhlE1*-L368L (2.4-fold), three of which also produced the highest nodulation gains.

To investigate whether improvements in nodulation competitiveness could be due to a better colonization of the plant culture medium or rhizosphere, we measured the survival of the 25 mutants in competition with their isogenic parental strain in these two compartments. Results from these assays showed that none of the adaptive mutations improved bacterial ability to colonize the culture medium and only 4 of them, including up^{115}-RSc0965, *efpR*-E66K and *rhlE1*-L368L, improved the colonization of the rhizosphere (S2.4 Fig) although by smaller factors than improvements in nodulation competitiveness. These results showed that improvements in nodulation competitiveness were generally not associated with a better colonization of the culture medium or of the rhizosphere, indicating that host entry is mainly controlled directly by the plant and is the dominant selective force driving the adaptation of legume symbionts in this evolution experiment.

Evolutionary modelling predicts that selection for host entry drives adaptation of legume symbionts

The preferential selection of nodulation competitiveness in our experiment prompted us to investigate the origin of this phenomenon. In particular, we wondered if stronger selection for host entry over within-host proliferation could be a general feature of symbiotic life cycles, or, instead, if it is more likely to be a specific property of our experimental system that may arise due to genetic constraints on symbiotic traits and/or specific features of the selective regime. We used computer simulations to model the evolution of bacterial populations that cycle

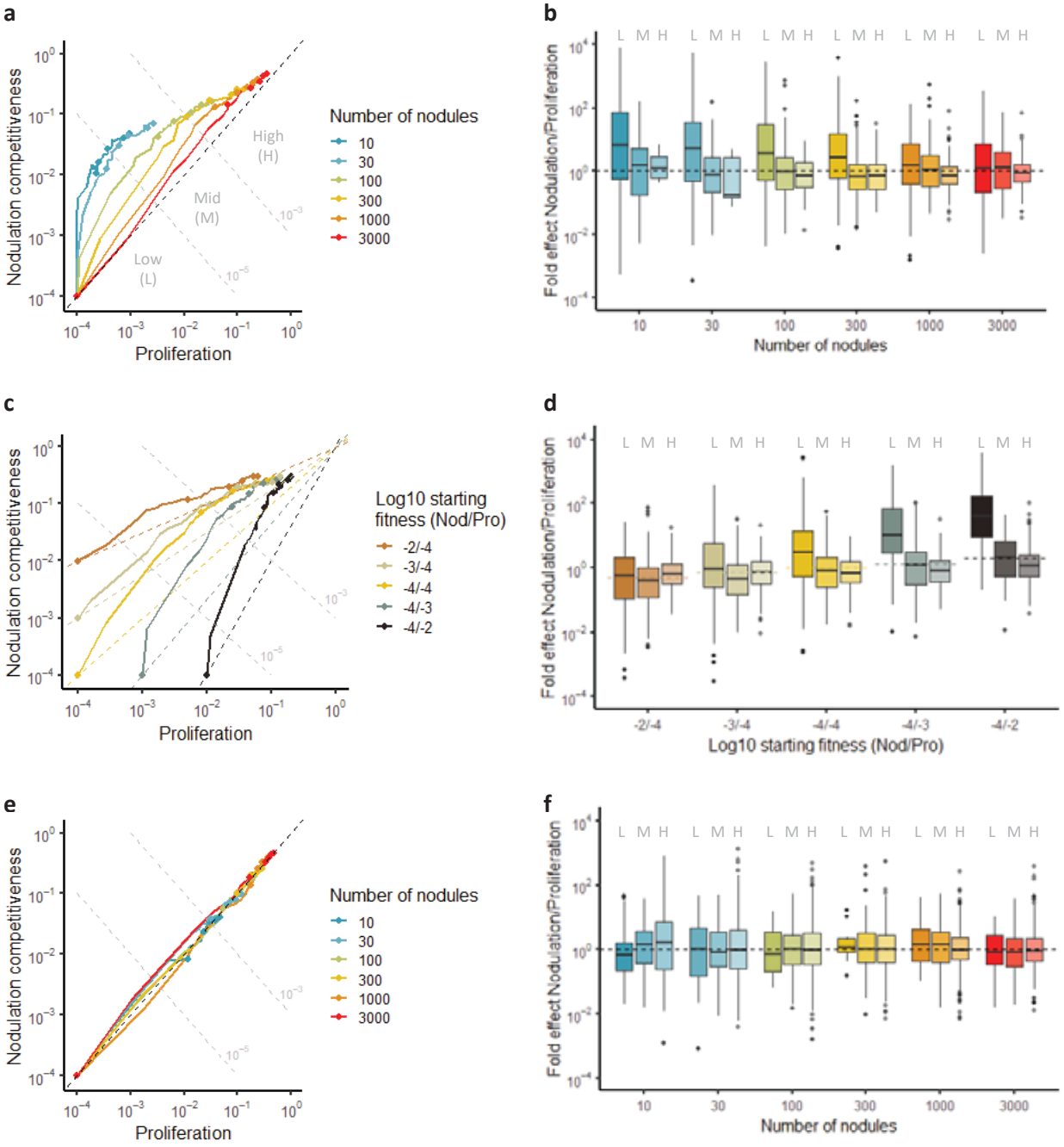


Fig 2.6. Relative strength of selection for nodulation competitiveness and proliferation

a,c,e, Median fitness trajectories of 100 simulated populations evolving under different sizes of nodulation bottleneck (10 to 3000 nodules). Points indicate fitness values at cycles 0, 10, 20, 30, 40 and 50. Dotted grey lines represent iso-fitness lines (defined as the product of nodulation competitiveness and proliferation) used to delineate 3 fitness domains: Low (L; fitness < 10^{-5}), Mid (M; 10^{-5} < fitness < 10^{-3}), and High (H; fitness > 10^{-3}). The black dotted line represents the diagonal, along which population would improve both phenotypic traits equally well. **b,d,f**, Fold effect (nodulation competitiveness effect divided by proliferation effect) of mutations that reached a frequency of at least 30% in simulated populations. Color shading indicate the fitness domain (Low, Mid or High) in which each group of mutations arose. Note that fold effects of mutations are indicated even when median fitness trajectories do not reach the High fitness domain (*e.g.* 10 and 30 nodule lines in **a**) because some individual replicate simulations (shown in S8-S16 Figs) do reach this fitness domain. The effect of different parameters on adaptive trajectories was tested: the size of nodulation bottleneck (**a, b**), the initial fitness values of the ancestor (**c, d**), and the size of nodulation bottleneck under a scenario where the nodulation bottleneck occurs after bacterial clonal proliferation (**e, f**). Raw data from evolutionary simulations are available at: https://doi.org/10.15454/QYB2S9.

between two compartments: the external environment (rhizosphere) and the host tissues (root nodules). In our model, bacteria accumulate mutations in the rhizosphere due to transient hypermutagenesis (Remigi et al., 2014) while we consider that the bacterial population remains clonal within nodules (Remigi et al., 2014) (see S2.5 Fig for full details on the model). These assumptions are relevant not only to rhizobium-legume interactions but also to other horizontally transmitted symbioses where the hosts are exposed to highly diverse environmental bacterial populations and accommodate more homogenous populations within their tissues (Fronk and Sachs, 2022; Kumar et al., 2015; Nyholm and McFall-Ngai, 2004). In our experimental system, bacterial fitness is mediated by two phenotypic components: competitiveness to enter the host (resulting in nodulation) and within-host proliferation. Host entry imposes a very stringent selective bottleneck since (i) in most cases, each nodule is founded by one single bacterium (Daubech et al., 2017; Gage, 2004), (ii) there were usually between 100 and 300 nodules per lineage collected at each cycle (S2.1 Table) while *ca*. 10^6-10^7 bacteria were inoculated at each cycle and (iii) bacterial genotypes differ in their competitive ability to form nodules (Fig 2.4a-b). Once inside the host, bacteria multiply to reach a carrying capacity that is directly proportional to their proliferationability, before returning to the external environment.

When following the evolution of bacterial phenotypes in populations founded by an ancestor with low initial fitness (10^{-4} fold compared to the theoretical optimum, for each phenotypic component), we observed, in most cases, a faster increase in nodulation competitiveness relative to within-host proliferation (Fig 2.6a). Accordingly, the distribution of mutations that were selected at the early steps of the adaptive process was biased towards stronger improvement of nodulation competitiveness over proliferation (Fig 2.6b, Low fitness domain). This dominance decreased, and was sometimes reverted, as populations progressed towards high fitness values. Decreasing the strength of the host entry bottleneck also reduced the dominance of early selection for nodulation competitiveness, until abolishing it for the largest bottleneck size tested (3000 nodules).

In our experiment, the ancestral nodulating strains had a much higher potential for improvement in nodulation competitiveness relative to proliferation ($5x10^4$ vs. 10^2-10^3 fold, respectively, compared to the natural symbiont *C. taiwanensis*) which could explain the faster evolution of the former trait. We performed additional simulations to test how evolutionary trajectories are affected by the genotype of the ancestor. As expected, starting from a higher

Table 2.1. Symbiotic phenotypes of beneficial mutations improving symbiotic fitness

Lineage	Gene ID[a]	Gene name	Mutation	Type of mutation[b]	Description	Cohort	Mean CI for *in planta* fitness[c]	Mean CI for plant culture medium colonization[d]	Mean CI for rhizosphere colonization[d]	Mean CI for nodulation competitiveness[c]	Mean gain factor in within-host proliferation[e]	Symbiotic phenotypes[f]	G score[g]	Adjusted *P* value[h]
B	RSp0595		P132P	Syn	Putative N-acyl-homoserine lactonase	CoB1	**77**	1	1	**5.2**	1.6	Nod+	2.80	1
	RSc1598		P455L	Nonsyn	Transmembrane sensor histidine kinase	CoB2	**24.2**	1	1.4	**4.7**	**6.4**	Nod+ Pro+	1.04	1
	RSp0788		E225K	Nonsyn	Carbamoyl-transferase	CoB3	**1.9**	1	1.1	**2**	1.2	Nod+	-0.09	1
	RSp1417		S417L	Nonsyn	Multidrug efflux transmembrane protein	CoB3	**1.6**	0.8	1.4	**3.9**	0.9	Nod+	12.47	**2.58E-04**
	RSp1590		S153L	Nonsyn	FAD-dependent oxireductase	CoB3	**6.3**	1	0.9	**2.4**	0.6	Nod+	4.10	1
	up[115]-RSc0965		G/A	Intergenic	Unknown function, negatively regulates *efpR* expression	CoB6	**31.1**	0.9	**2.9**	**8.2**	**2.6**	Rhizo+ Nod+ Pro+	nd	nd
	RSc0539	*rhlE1*	L368L	Syn	ATP-dependent RNA helicase protein	CoB7	**46.2**	1.1	**4**	**28.2**	**2.4**	Rhizo+ Nod+ Pro+	2.88	1
	RSc0243		L213L	Syn	LysR-type transcription regulator protein	CoB10	**3**	1.3	1.4	**3.9**	1.1	Rhizo+ Nod+	1.16	1
	RSp1116		I32I L33F	Syn Nonsyn	Putative low specificity L-threonine aldolase	CoB10	**2.1**	1.2	1.2	**2.1**	0.9	Nod+	0.84	1
	up[8]-RSp1469		C/T	Intergenic	Putative tyrosine specific protein phosphatase	CoB10	**2.1**	0.9	1	**2**	1.0	Nod+	nd	nd
	pRALTA_0230		G242E	Nonsyn	Putative ATP-dependent DNA helicase	CoB10	**1.5**	1.5	1.5	**2.1**	0.9	Nod+	6.34	1
	pRALTA_0552		I100I	Syn	Pseudogene, putative membrane bound hydrogenase	CoB10	**2.9**	1.2	1.4	**2.3**	0.8	Nod+	nd	nd
G	RSc1097	*efpR*	E66K	Nonsyn	Transcriptional regulator	CoG1	**212.1**	1.3	**4**	**35**	**3.7**	Rhizo+ Nod+ Pro+	3.17	1
	RSc1982		I100I	Syn	DNA methyltransferase	CoG2	**3.3**	3.1	1.3	**4.0**	0.7	Nod+	1.44	1
	RSc1164	*ppiB*	Q112*	Nonsense	Peptyl-prolyl-cis/trans isomerase	CoG4	**24.5**	1	1.0	**5.0**	1.4	Nod+	2.26	1
	RSc2277		V321G	Nonsyn	Putative transporter protein	CoG4	**133.9**	9.3	10.1	**58**	**2.5**	Nod+ Pro+	36.95	**4.68E-50**
	RSc0194		S102F	Nonsyn	Zinc-dependent alcohol dehydrogenase	CoG13	**5.5**	1.1	1.6	**3.1**	1.4	Nod+	4.67	1
	RSc1633	*phbB*	S38F	Nonsyn	Acetoacetyl-CoA reductase	CoG13	**7.9**	2	2.1	**3.3**	**1.7**	Rhizo+ Nod+ Pro+	1.62	1
	RSp0741	*tssJ*	I157N	Nonsyn	Type VI secretion system, lipoprotein TssJ	CoG13	**6.7**	1	1.1	**2.9**	1.7	Nod+	6.99	1
	RSp0574		L15L	Syn	Hypothetical protein	CoG15	**3.2**	1.1	1.3	**2.1**	**1.6**	Nod+ Pro+	4.36	1
	RSp0587		P84P	Syn	Putative signal peptide hypothetical protein	CoG15	**2**	1.1	0.8	**2.6**	1.3	Nod+	28.74	**1.04E-46**
	RSp0759		F91F P92S	Syn Nonsyn	Putative type VI secretion-associated protein	CoG15	**5.4**	0.7	0.9	**2.5**	1.4	Nod+	1.00	1
	RSp0769		F416S	Nonsyn	Putative type VI secretion-associated protein	CoG15	**6.3**	1.1	1.3	**1.6**	**1.8**	Nod+ Pro+	17.20	**3.98E-10**
	up[116]-RSp0780		C/T	Intergenic	Hypothetical protein	CoG15	**2.5**	1.1	1.1	**3**	1.1	Nod+	nd	nd
	RSp1422		G1298E	Nonsyn	Non-ribosomal peptide synthase	CoG15	**2**	1.2	1.2	**2.4**	**1.6**	Nod+	18.28	**2.79E-03**

[a]up[x], intergenic mutations located x nucleotides upstream the gene indicated. [b]Syn, synonymous mutations. Nonsyn, non synonymous mutations. [c]Mean values of competitive indexes (CI) obtained from 3 to 4 independent experiments. Figures in bold are statistically different from 1 (P <0.05, t -test with the Benjamini-Hochberg correction). dMean values obtained from at least 15 measurements from 3 to 4 independent experiments. Figures in bold are statistically significant (P <0.05, *Wilcoxon* test with the Benjamini-Hochberg correction). [e]Mean values obtained from at least 15 measurements from 3 to 4 independent experiments. Figures in bold are statistically different from 1 (P <0.05, Wilcoxon test with Benjamini-Hochberg correction). [f]Rhizo+, improvement in rhizosphere colonization. Nod+, improvement in nodulation competitiveness. Pro+, improvement in within-host proliferation. [g]G scores were calculated as proposed by Tenaillon et al. [23]. nd, G scores were not determined for mutations in pseudogenes or intergenic regions. [h]P-values associated with G scores were adjusted using a Bonferroni correction to evaluate genetic parallelism.

proliferation fitness in the ancestor further increased the overall fold effect of early selected mutations on nodulation competitiveness (Fig 2.6c-d). In contrast, starting with a higher initial nodulation competitiveness value of 10^{-3} or 10^{-2} in the ancestor was not sufficient to lead to a stronger selection for proliferation, further confirming the asymmetry between selective pressures acting on the two phenotypes. These findings were qualitatively robust to our assumptions regarding the genetic architecture of the two symbiotic phenotypes (S2.6-S2.15 Figs).

Finally, we wondered if the chronology of symbiotic steps impacts on selective pressures and evolutionary trajectories. We modified the simulations so that clonal multiplication of bacteria (according to their proliferation fitness) occurs before the selective bottleneck. Although this model is not relevant in the context of the legume-rhizobium symbiosis, it may apply to other symbiotic or pathogenic life cycles, for example when a selective bottleneck occurs during dissemination in the environment or within the host (Hullahalli and Waldor, 2021; Joseph et al., 2015), as well as in the case of vector-mediated transmitted pathogens (Goodrich-Blair and Clarke, 2007; Koinuma et al., 2020) or vertically transmitted symbionts (Perreau et al., 2021). In this context, the dominance of nodulation competitiveness was strongly reduced, bringing the two selective forces close to equilibrium (Fig 2.6e-f and S2.16 Fig), in agreement with the expectation that each fitness component should have a similar influence on adaptation when they have a multiplicative effect on global fitness. Therefore, the dominance of nodulation competitiveness in the previous simulation runs is likely a consequence of the selective bottleneck occurring before within-host bacterial multiplication and just after the transient hypermutagenesis phase in the rhizosphere.

Altogether, our simulations show that the selective pressures that are experienced by symbiotic bacteria are asymmetric, with a dominance of selection for nodulation competitiveness over within-host proliferation. This asymmetry occurs when a genetically and phenotypically diverse bacterial population from the rhizosphere is exposed to the strong selective bottleneck at host entry, allowing the efficient selection of the most competitive clones by the host plant. The fact that the ancestor from our evolution experiment has a lower nodulation competitiveness value compared to proliferation (relative to that of the wild-type rhizobium *C. taiwanensis*) is an additional factor expected to have contributed to the stronger selection on nodulation competitiveness. Importantly, the key factors responsible for the dominance of selection for host entry competitiveness in our model (strong bottleneck at host

entry and phenotypic diversity of environmental microbial populations) are found in other symbiotic systems, suggesting that competition for host-entry is probably an important driver of bacterial adaptation in emerging symbiotic associations.

Discussion

Horizontally transmitted symbiotic bacteria have complex lifecycles during which they have to face multiple environmental constraints, both outside and inside the host. Assessing the relative influence of each of these selective pressures on the adaptive trajectories of emerging symbiotic bacteria is critically needed to better understand the eco-evolutionary dynamics of microbial populations (Obeng et al., 2021).

Here we show that selection on competitiveness for host entry is the dominant selective pressure shaping early adaptation of new rhizobia. This trait, mainly mediated by the plant, improved very fast during our evolution experiment, to a level comparable to that of the natural *Mimosa* symbiont *C. taiwanensis*, and all adaptive mutations identified in lineages B and G improved nodulation competitiveness. Evolutionary simulations led us to propose that the dominance of the selection on host entry over the selection on within-host proliferation that we observed in our system is strongly dependent on the strength of the selective bottleneck that occurs at host entry and its position in the life cycle of symbionts. These results shed light on the importance of selective bottlenecks in shaping the evolutionary trajectories of emerging legume symbionts. Most theoretical (Bergstrom et al., 1999; LeClair and Wahl, 2018) and experimental (Garoff et al., 2020; Izutsu et al., 2022; Mahrt et al., 2021; Sousa et al., 2017; Wein and Dagan, 2019; Windels et al., 2021) works on the influence of bottlenecks on microbial adaptation have focused on non-selective bottlenecks that randomly purge genetic diversity and reduce the efficiency of natural selection. Yet, another aspect of bottlenecks emerges when considering that transmission and host colonization can be, at least partially, dependent on microbial genotype, prompting us to consider infection bottlenecks as selective events (Handel & Bennett, 2008). Selective bottlenecks have already been described during virus infections (Huang et al., 2017; Joseph et al., 2015; Moncla et al., 2016; Sobel Leonard et al., 2016; Wilker et al., 2013) but received less attention in bacteria, except for their role in the evolution of phenotypic heterogeneity (De Ste Croix et al., 2020; Libby and Rainey, 2011; Moxon and Kussell, 2017; Schmutzer and Wagner, 2020). Yet, potentially all host-associated bacterial populations (whether commensal, pathogenic or mutualistic)

experience bottlenecks of various intensities (sometimes going down to only a few cells (Gama et al., 2012) during their life cycles (Abel et al., 2015; Fronk and Sachs, 2022; Papkou and A Gokhale, 2016). Since competitiveness for host entry is a complex, polygenic phenotypic trait (Bright and Bulgheresi, 2010; Cossart and Sansonetti, 2004; Rowe et al., 2019; Visick et al., 2021; Wheatley et al., 2020) that often displays extensive variation within natural populations (Bongrand et al., 2020; Mendoza-Suárez et al., 2021), bottlenecks are expected to represent major selective events (instead of purely stochastic events) in many symbiotic life cycles. Interestingly, strong selection for host entry in experimentally evolved mutualistic and pathogenic symbionts was recently observed in two studies (Bacigalupe et al., 2019; Robinson et al., 2018) reinforcing the idea that selective bottlenecks at host entry may play a crucial role in the evolution of diverse host-microbe interactions.

In multistep infection processes, the different infection steps can be functionally linked either by couplings or trade-offs (Ebert et al., 2016; Hall et al., 2019). Our study unveiled a new characteristic of the rhizobium-legume interactions: the recurrent genetic coupling between nodulation competitiveness and within-host proliferation. Indeed, all the 9 identified mutations that improved within-host proliferation also improved nodulation competiveness, while the reverse was not true. This result corroborates our previous data showing that mutations promoting nodule cell infection always improved nodule formation (Guan et al., 2013; Marchetti et al., 2010). It is conceivable that the intracellular release and proliferation of rhizobia in nodule cells is dependent on the efficiency of the earliest symbiotic events, *i. e*. the entry and progression of bacteria in infection threads and the concomitant divisions of nodule cells preparing for the accommodation of bacteria. Consistently, delayed progression of infection threads across the root cell layers was shown to impair the release of bacteria in nodule cells (Xiao et al., 2014). Moreover the identification of plant receptors (LYK3, NFR1, Sym37, SYMRK) and transcription factors (NIN, ERN1, ERN2) involved in both nodule organogenesis and rhizobial intracellular infection (Capoen et al., 2005; Cerri et al., 2016; Liu et al., 2021; Moling et al., 2014) supports the existence of common mechanisms controlling the two processes on the plant side. Future functional analyses of the adaptive mutations identified in this study will expand our understanding of the molecular bases of nodulation competitiveness in rhizobia (Bovin & Lepetit, 2020; Mendoza-Suárez et al., 2021; Younginger & Friesen, 2019), and its relationship to within-host proliferation. Including mutualistic traits (nitrogen fixation and host growth promotion) in the analysis of the genetic couplings (or

trade-offs) between the different symbiotic traits will be important to fully characterize the genetic constraints shaping the evolution of rhizobium-legume interactions (Batstone et al., 2020; Quides et al., 2021; Younginger and Friesen, 2019). The genetic links between different phases of host colonization are generally poorly documented in other host-microbe associations, but they were analyzed in recent experimental evolution studies with opposite outcomes. Robinson et al. (2018) showed that early bacterial adaptation to zebrafish gut favored the improvement of host entry and inter-host transmission without affecting within-host proliferation (Robinson et al., 2018). In another study, experimental evolution of *Vibrio fisherii* in symbiosis with squids led to the fixation of mutations in a global regulatory gene (*binK*) that improved both the initiation and maintenance of the interaction, through its action on several bacterial phenotypic traits (Pankey et al., 2017). Conversely, an example of trade-off between early and late infection stages was recently evidenced in *Xenorhabdus nematophila* interacting with insects (Cambon et al., 2019; Faucher et al., 2021). In the same line, a trade-off between within-vector and within-host fitness was evidenced in the mosquito-borne parasite *Plasmodium falciparum* causing malaria (Costa et al., 2018). These examples highlight the diversity and context-dependence of genetic correlations that exist between bacterial phenotypic traits involved in symbiotic interactions.

In all lineages of this experiment, the rate of adaptation was very high during the first cycles of evolution and then tended to decrease over time. This classical pattern of evolution, due to diminishing return epistasis among beneficial mutations (Couce and Tenaillon, 2015; McDonald, 2019; Van den Bergh et al., 2018), was likely favored by the low initial fitness values of nodulating ancestors compared to the natural symbiont *C. taiwanensis* allowing the rapid acquisition of highly beneficial mutations during the first cycles (Table 1). Elevated mutation rate in the rhizosphere, prior to the entry of bacteria in roots (Remigi et al., 2014), has also played an important role in the dynamic of adaptation by fueling the extensive genetic diversification of bacteria and exposing these populations to plant-mediated selection. A probable consequence of high mutation rate combined to strong selection is the presence of large cohorts of mutations that carry multiple beneficial mutations along with some neutral or slightly deleterious ones. Mutational cohorts were already described in previous evolution experiments (Lang et al., 2013; Maddamsetti et al., 2015), with cases of cohorts carrying 2 co-driver mutations (Buskirk et al., 2017). Here we identified up to 6 adaptive mutations in one of our cohorts. The occurrence of multi-driver mutations per cohort might be explained by the

nested, sequential fixation of multiple adaptive mutations in one lineage before it reaches detectable frequency (5% in our case) (Buskirk et al., 2017). This phenomenon, creating a 'travelling wave' of adaptation (Neher, 2013), was recently observed with high-resolution sequencing (Nguyen Ba et al., 2019) and is possibly amplified by the strong selection/high mutation regime of our experiment. Alternatively, it was proposed that, in the case of hypermutagenesis, the likelihood of multiple adaptive mutations arising simultaneously in a given genome becomes non-negligible and might promote saltational evolution (Katsnelson et al., 2019).

In conclusion, our work shows that the relative strengths of the selective forces imposed by legumes strongly influence the evolutionary trajectory of symbiotic microbial populations. Selection for host entry is likely the predominant life-history trait selected during the evolution of new rhizobia. This effect emerges from the presence of a selective bottleneck at host entry, although its predominance can be modulated by other genetic and ecological factors. Investigating the role of selective bottlenecks in other symbiotic interactions will improve our ability to manage and predict the eco-evolutionary dynamics of host-associated microbial populations.

Methods

Bacterial strains and growth conditions

Strains and plasmids used in this study are listed in the S2.6 Table. *Cupriavidus taiwanensis* strains were grown at 28°C on tryptone-yeast (TY) medium (Beringer). *Ralstonia solanacearum* strains were grown at 28°C either on BG medium (Boucher et al., 1985) or on minimal MP medium (Plener et al., 2010) supplemented with 2% glycerol. *Escherichia coli* strains were grown at 37°C on Lysogeny-Broth medium (BERTANI, 1951). Antibiotics were used at the following concentrations: trimethoprim at 100 µg ml^{-1}, spectinomycin at 40 µg ml^{-1}, streptomycin at 200 µg ml^{-1} and kanamycin at 25 µg ml^{-1} (for *E. coli*) or 50 µg ml^{-1} (for *R. solanacearum*).

Experimental evolution

Symbiotically evolved clones and populations were generated as previously described (Marchetti et al., 2017). Five lineages, two (B and F) derived from the CBM212 ancestor, two (G and K) derived from the CBM349 ancestor and one (M) derived from the CBM356 ancestor,

previously evolved for 16 cycles were further evolved until cycle 35. For each lineage, at each cycle, 30 plants were inoculated with nodule bacterial populations from the previous cycle. Twenty-one days after inoculation, all nodules from the 30 plants were pooled, surface-sterilized with 2.6% sodium hypochlorite for 15 min, rinsed with sterile water and crushed. Then, 10% of the nodule crush was used to inoculate a new set of 30 plants the same day. Serial dilutions of each nodule crush were plated and one clone was randomly selected from the highest dilution and purified. The selected clones, the rest of the nodule crushes, and a 24-h culture of bacteria from an aliquot of nodule crushes were stored at –80°C at each cycle.

Sequencing bacterial populations and clones

Aliquots of frozen nodule bacterial populations or single colonies from purified clones were grown overnight in BG medium supplemented with trimethoprim. Bacterial DNA was extracted from 1 ml of culture using the Wizard genomic DNA purification kit (Promega). Evolved population DNAs were sequenced at the GeT-PlaGe core facility (https://get.genotoul.fr/), INRAE Toulouse. DNA-seq libraries have been prepared according to Illumina's protocols using the Illumina TruSeq Nano DNA HT Library Prep Kit. Briefly, DNA was fragmented by sonication, size selection was performed using SPB beads (kit beads) and adapters were ligated to be sequenced. Library quality was assessed using an Advanced Analytical Fragment Analyzer and libraries were quantified by qPCR using the Kapa Library Quantification Kit (Roche). Sequencing has been performed on a NovaSeq6000 S4 lane (Illumina, California, USA) using a paired-end read length of 2x150 pb with the Illumina NovaSeq Reagent Kits.

Evolved clones were re-sequenced either by C.E.A/IG/Genoscope using the Illumina GA2X technology (clones B8, B16, F16, G8, G16, K8, K16, M8 and M16), or by the GeT_PlaGe core facility using the Illumina technology HiSeq2000 (clones B1, B2, B3, B4, B5, B6, B7, B9, B10, B11, B12, B13, B14, B15, G1, G2, G3, G4, G5, G6, G7, G9, G10, G11, G12, G13, G14, G15, M4 and M12), or HiSeq3000 (clones B20, B25, B30, G25, G30, K1, K2, K3, K4, K5, K6, K9, K10, K11, K12, K13, K14, K15, M2, M6, M10 and M14), or MiSeq (K7) or NovaSeq6000 (clones B35, F35, G35, K35 and M35).

Detection of mutations and molecular analyses

Sequencing reads from NovaSeq6000 runs (whole-populations and clones from cycle 35) were mapped on the chimeric reference genome of the ancestral strain, comprising *R. solanacearum* GMI1000 chromosome (GenBank accession number: NC_003295.1) and megaplasmid (NC_003296.1) together with *C. taiwanensis* symbiotic plasmid pRalta (CU633751). Mutations were detected using breseq v0.33.1 (Deatherage & Barrick, 2014) with default parameters, either using the polymorphism mode (for whole-population sequences) or the consensus mode (for individual clones). Mutation lists were curated manually in order to remove mutations present in the ancestral strains as well as false positive hits arising from reads misalignments in low complexity regions. Moreover, alleles showing aberrant trajectories in the time-course whole-population sequencing data were checked manually, by inspecting either breseq output files and/or read alignments with IGV (Robinson et al., 2011), and corrected as needed. In this work, we focused our attention on SNPs and indels detected above 5% frequency in the populations. Recombinations, rearrangements and IS movements were not analyzed exhaustively.

Sequencing data of the remaining clones were analyzed as described previously (Capela et al., 2017) with the PALOMA bioinformatics pipeline implemented in the Microscope platform (Vallenet et al., 2020). The complete list of mutational events generated for these clones are available on the Microscope platform (https://mage.genoscope.cns.fr/microscope/expdata/NGSProjectEvo.php, SYMPA tag).

In order to identify mutations forming temporal clusters (cohorts) through populations, we selected mutations having a frequency of 30% in at least one cycle and clustered their frequency using the *hclust* function in R v3.6.1 (Team, 2008). Then, we separated cohorts using a cutoff distance of 0.3.

Sub-population genealogies in lineages B and G, shown as Muller plots in S2.3 Fig, were reconstructed by comparing mutations found in cohorts and in individually sequenced clones. Cohorts for which ancestry cannot be ascertained (mutations not found in any clone) were not included in these plots. Relative frequencies of genotypes were calculated manually and plotted with the R package MullerPlot (Noble, 2019).

G scores were calculated as described (Tenaillon et al., 2016). Since we found that synonymous mutations can be adaptive, we used all types of mutations in coding regions to calculate this statistics. Moreover, we included all mutations beyond 5% frequency in this

analysis since we assumed that strong clonal interference may prevent adaptive mutations to rise to high frequency. To evaluate statistical significance of G scores, we ran 1,000 simulations where the total number of mutations used for the calculation of the observed G scores (3330) were randomly attributed in the coding genes according to their respective length to compute simulated G scores for each bacterial gene. The sum of simulated G scores was compared to the observed sum and we computed a Z score and *P*-value from these simulated G statistics. These simulations were also used to compute mean G scores for each gene, and to calculate the associated Z scores and *P*-values (adjusted using a Bonferroni correction).

Reconstruction of mutations in evolved clones

Mutant alleles and constitutively expressed reporter genes (GFPuv, mCherry) were introduced into *Ralstonia* evolved clones by co-transformation using the MuGent technique (Capela et al., 2017; Dalia et al., 2014). Briefly, two DNA fragments, the first one carrying an antibiotic (kanamycin) resistance gene prepared from the pRCK-P*ps*-GFP or pRCK-P*ps*-mCherry linearized plasmids allowing the integration of the resistance gene in the intergenic region downstream the *glmS* gene, and the second one carrying the mutation to be introduced prepared by PCR amplification of a 6 kb region using genomic DNA of evolved clones as template and high fidelity Phusion polymerase (ThermoFisher Scientific), were co-transformed into naturally competents cells of *Ralstonia* evolved clones. Co-transformants resistant to kanamycin were screened by PCR using primers specifically amplifying mutant or wild-type alleles and verified by Sanger sequencing. Primers used in mutation reconstructions are listed in S6 Table.

Analyses of symbiotic phenotypes on *M. pudica*

Mimosa pudica seeds (LIPME 2019 production obtained from one commercial seed (BandT World Seed, Paguignan, France) of Australian origin) were sterilized as described (Chen, Moulin, et al., 2003) by immersion in 95% H_2SO_4 during 15 minutes and 2.4% sodium hypochlorite solution during 10 minutes and rinsed in sterile distilled water 5 times. Seeds were soaked in sterile water at 28°C under agitation for 24 hours and then deposited on soft agar (9.375 g/L) and incubated at 28°C during 24 more hours in darkness. Then, seedlings were cultivated in glass tubes (2 seedlings/tube) under N-free conditions, each tube containing 20 mL of solid Fahraeus medium (Fahraeus, 1957) and 40 mL of liquid Jensen medium (Jensen,

1942) diluted 1:4 with sterile water. Plants were incubated in a culture chamber at 28°C under a photoperiod day/night of 16 h/8 h and positioned randomly in the different experiments.

For symbiotic relative fitness and nodulation competitiveness assays, two strains of *Ralstonia solanacearum* expressing differential constitutive fluorophores (GFPuv or mCherry) or one strain of *R. solanacearum* and one strain of *C. taiwanensis*, both expressing differential antibiotic resistance (streptomycin and trimethoprim), were co-inoculated onto *M. pudica* plantlets grown for 3 to 4 days in the culture chamber. Both strains were inoculated in equivalent proportion (*ca*. $5x10^5$ bacteria of each strain per plant) except for the comparisons of the *Ralstonia* nodulating ancestors with *C. taiwanensis* where strains were mixed in a 1000:1 ratio (*ca.* $5x10^3$ bacteria of *C. taiwanensis* and *ca.* $5x10^5$ bacteria of nodulating ancestors per plant). Nodules were harvested 21 days after inoculation, surface sterilized by immersion in 2.4% sodium hypochlorite solution for 15 minutes and rinsed with sterile water. For symbiotic relative fitness measurements, sterilized nodules from 20 plants were pooled and crushed in 1 mL of sterile water, diluted and spread on selective solid medium using an easySpiral automatic plater (Interscience). After 2-day incubation at 28°C, colonies were screened either by plating on selective medium in case of co-inoculations of *Ralstonia* evolved clones with *C. taiwanensis* or based on fluorescence using a stereo zoom microscope (AxioZoom V16, Zeiss, Oberkochen, Germany) for co-inoculations of *Ralstonia* evolved clones with *Ralstonia* reconstructed mutants. For nodulation competitiveness assays, *ca*. 96 individual nodules per experiment were crushed separately in 96-well microtiter plates and droplets were deposited on selective medium. Nodule occupancy was determined by screening bacteria grown in the droplets either on selective medium or based on their fluorescence as described for fitness measurements. Both nodulation competitiveness and relative fitness assays were measured in at least three independent biological replicates. Competitive indexes (CI) were calculated by dividing the ratio of the number of test strain (evolved strains or reconstructed mutants) vs. reference strain (*C. taiwanensis* or evolved parental strains, respectively) in nodules normalized by the inoculum ratio. When CI values were all above 1, CI values were transformed by their inverse and compared to the value 1 using a one-sided Student *t*-test with Benjamini-Hochberg correction ($P<0.05$).

For within-host proliferation assays, plantlets were inoculated with a single strain ($5x10^5$ bacteria per plant). In each experiment, nodules from 6 individual plants were collected separately 21 days after inoculation, surface sterilized for 15 minutes in a 2.4% sodium

hypochlorite solution, rinsed with sterile water and crushed in 1 ml of sterile water. Dilutions of nodule crushes were plated on selective solid medium using an easySpiral automatic plater (Interscience). Two days after incubation at 28°C, the number of colonies was counted. Within-host proliferation was estimated as the number of bacteria per nodule. For each strain, 15 to 24 measurements from three independent biological replicates were performed. Pairwise comparisons of proliferation values were compared using a two-sided Wilcoxon rank sum test with Benjamini-Hochberg correction.

For Jensen culture medium and rhizosphere colonization assays, plantlets were co-inoculated with pair of strains expressing differential constitutive fluorophores (GFPuv or mCherry) in equivalent proportion ($5x10^6$ bacteria of each strain per plant). Seven days after inoculation, bacteria present in the culture medium were diluted and plated on selective solid medium using an easySpiral automatic plater (Interscience). Bacteria attached to the roots were resuspended in 4 ml of sterile water by strongly vortexing for 1 minute, diluting and plating on selective medium using an easySpiral automatic plater (Interscience). Two days after incubation at 28°C, colonies were screened for fluorescence using a stereo zoom microscope (AxioZoom V16, Zeiss, Oberkochen, Germany). Competitive indexes (CI) for Jensen medium or rhizosphere colonization were calculated as described above. CI values were compared to the value 1 using a two-sided Wilcoxon rank sum test with Benjamini-Hochberg correction.

Sample sizes were determined based on our previous studies (Capela et al., 2017; Guan et al., 2013; Tang et al., 2020). In cases where we performed only three independent replicates, each replicate was based on the analysis of a large number of plants or nodules. Independent biological replicates were performed on different days, with different plants and different bacterial cultures.

Modelling/simulations

Full details on evolutionary simulations and the list and values of parameters used in the simulations are available S2.7 Table.

Data availability statement

Sequencing data are available under NCBI SRA BioProject ID PRJNA788708 and SRP353965. Raw experimental data are available as Supplementary Tables. Raw data generated from

computer simulations are deposited on the Data INRAE dataverse (https://doi.org/10.15454/QYB2S9).

Code availability statement

Computer code used to analyze genomic data and to perform computer simulations are deposited on the Data INRAE dataverse (https://doi.org/10.15454/QYB2S9).

Acknowledgements

We are grateful to A. Carlier and F. Roux for careful reading of the manuscript, to O. Tenaillon for helpful discussions on the project, to O. Tenaillon and H. Le Nagard for sharing their genetic algorithm script, and to Ludovic Legrand for archiving sequencing data and their associated metadata. We acknowledge the Genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, https://doi.org/10.15454/1.5572369328961167E12) for providing computing and storage resources.

G.G.D.d.M was supported by a fellowship from the French Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI). P.R. received funding from the EU in the framework of the Marie-Curie FP7 COFUND People Programme, through the award of an AgreenSkills+ fellowship (under grant agreement n°609398) and from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N° 845838.

This study was supported by the Fédération de Recherche Agrobiosciences, Interactions et Biodiversité, the French National Research Agency (ANR-16-CE20-0011-01 and ANR-21-CE02-0019-01) and the "Laboratoires d'Excellence (LABEX)" TULIP (ANR-10-LABX-41)" and the "École Universitaire de Recherche (EUR)" TULIP-GS (ANR-18-EURE-0019).

Author contributions

Conceptualization: Catherine Masson-Boivin, Delphine Capela, Philippe Remigi

Data curation: Ginaini Grazielli Doin de Moura, Delphine Capela, Philippe Remigi

Formal analysis: Ginaini Grazielli Doin de Moura, Delphine Capela, Philippe Remigi

Funding acquisition: Catherine Boivin-Masson, Philippe Remigi, Jean-Baptiste Ferdy

Experiments: Ginaini Grazielli Doin de Moura, Saida Mouffok, Nil Gaudu, Anne-Claire Cazalé, Marine Milhes, Tabatha Bulach, Sophie Valière, David Roche, Delphine Capela, Philippe Remigi

Supervision: Delphine Capela, Philippe Remigi

Writing – original draft: Ginaini Doin de Moura, Delphine Capela, Philippe Remigi

Writing – review and editing: Catherine Masson-Boivin, Jean-Baptiste Ferdy, Delphine Capela, Philippe Remigi.

Competing interests

The authors declare no competing interests.

References

1. Drew GC, Stevens EJ, King KC. Microbial evolution and transitions along the parasite-mutualist continuum. Nat Rev Microbiol. 2021; 19(10):623-38. doi: 10.1038/s41579-021-00550-7. PMID: 33875863.
2. Nandasena KG, O'Hara GW, Tiwari RP, Sezmis E, Howieson JG. In situ lateral transfer of symbiosis islands results in rapid evolution of diverse competitive strains of mesorhizobia suboptimal in symbiotic nitrogen fixation on the pasture legume *Biserrula pelecinus* L. Environ Microbiol. 2007; 9(10):2496-511. doi: 10.1111/j.1462-2920.2007.01368.x. PMID: 17803775.
3. Longdon B, Hadfield JD, Day JP, Smith SC, McGonigle JE, Cogni R, et al. The causes and consequences of changes in virulence following pathogen host shifts. PLoS Pathog. 2015; 11(3):e1004728. doi: 10.1371/journal.ppat.1004728. PMID: 25774803.
4. Bonneaud C, Longdon B. Emerging pathogen evolution: Using evolutionary theory to understand the fate of novel infectious pathogens. EMBO Rep. 2020; 21(9):e51374. doi: 10.15252/embr.202051374. PMID: 32864788.
5. Obeng N, Bansept F, Sieber M, Traulsen A, Schulenburg H. Evolution of Microbiota-Host Associations: The Microbe's Perspective. Trends Microbiol. 2021; 29(9):779-87. doi: 10.1016/j.tim.2021.02.005. PMID: 33674142.
6. Robinson CD, Bohannan BJ, Britton RA. Scales of persistence: transmission and the microbiome. Curr Opin Microbiol. 2019; 50:42-9. doi: 10.1016/j.mib.2019.09.009. PMID: 31629296.

7. Wheatley RM, Ford BL, Li L, Aroney STN, Knights HE, Ledermann R, et al. Lifestyle adaptations of Rhizobium from rhizosphere to symbiosis. Proc Natl Acad Sci U S A. 2020; 117(38):23823-34. doi: 10.1073/pnas.2009094117. PMID: 32900931.
8. Gage DJ. Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*. J Bacteriol. 2002; 184(24):7042-6. doi: 10.1128/jb.184.24.7042-7046.2002. PMID: 12446653.
9. Gage DJ. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol Mol Biol Rev. 2004; 68(2):280-300. doi: 10.1128/mmbr.68.2.280-300.2004. PMID: 15187185.
10. Coba de la Peña T, Fedorova E, Pueyo JJ, Lucas MM. The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organelle? Front Plant Sci. 2017; 8:2229. doi: 10.3389/fpls.2017.02229. PMID: 29403508.
11. Hirsch AM, Wilson KJ, Jones JD, Bang M, Walker VV, Ausubel FM. *Rhizobium meliloti* nodulation genes allow *Agrobacterium tumefaciens* and *Escherichia coli* to form pseudonodules on alfalfa. J Bacteriol. 1984; 158(3):1133-43. PMID: 6327629.
12. Abe M, Kawamura R, Higashi S, Mori S, Shibata M, Uchiumi T. Transfer of the symbiotic plasmid from *Rhizobium leguminosarum* biovar trifolii to *Agrobacterium tumefaciens*. J Gen Appl Microbiol. 1998; 44(1):65-74. PMID: 12501295.
13. Nakatsukasa H, Uchiumi T, Kucho K, Suzuki A, Higashi S, Abe M. Transposon mediation allows a symbiotic plasmid of *Rhizobium leguminosarum* bv. trifolii to become a symbiosis island in *Agrobacterium* and *Rhizobium*. J Gen Appl Microbiol. 2008; 54(2):107-18. doi: 10.2323/jgam.54.107. PMID: 18497485.
14. Marchetti M, Capela D, Glew M, Cruveiller S, Chane-Woon-Ming B, Gris C, et al. Experimental evolution of a plant pathogen into a legume symbiont. PLoS Biol. 2010; 8(1). doi: e100028010.1371/journal.pbio.1000280. PMID: 20084095.
15. Masson-Boivin C, Giraud E, Perret X, Batut J. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? Trends Microbiol. 2009; 17(10):458-66. doi: 10.1016/j.tim.2009.07.004. PMID: 19766492.
16. Marchetti M, Clerissi C, Yousfi Y, Gris C, Bouchez O, Rocha E, et al. Experimental evolution of rhizobia may lead to either extra- or intracellular symbiotic adaptation depending on the selection regime. Mol Ecol. 2017; 26(7):1818-31. doi: 10.1111/mec.13895. PMID: 27770459.

17. Guan SH, Gris C, Cruveiller S, Pouzet C, Tasse L, Leru A, et al. Experimental evolution of nodule intracellular infection in legume symbionts. ISME J. 2013; 7(7):1367-77. doi: 10.1038/ismej.2013.24. PMID: 23426010.
18. Capela D, Marchetti M, Clerissi C, Perrier A, Guetta D, Gris C, et al. Recruitment of a lineage-specific virulence regulatory pathway promotes intracellular infection by a plant pathogen experimentally evolved into a legume symbiont. Mol Biol Evol. 2017; 34(10):2503-21. doi: 10.1093/molbev/msx165. PMID: 28535261.
19. Tang M, Bouchez O, Cruveiller S, Masson-Boivin C, Capela D. Modulation of quorum sensing as an adaptation to nodule cell infection during experimental evolution of legume symbionts. mBio. 2020; 11(1). doi: 10.1128/mBio.03129-19. PMID: 31992622.
20. Moxon R, Kussell E. The impact of bottlenecks on microbial survival, adaptation, and phenotypic switching in host-pathogen interactions. Evolution. 2017; 71(12):2803-16. doi: 10.1111/evo.13370. PMID: 28983912.
21. Remigi P, Capela D, Clerissi C, Tasse L, Torchet R, Bouchez O, et al. Transient hypermutagenesis accelerates the evolution of legume endosymbionts following horizontal gene transfer. PLoS Biol. 2014; 12(9):e1001942. doi: 10.1371/journal.pbio.1001942. PMID: 25181317.
22. Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, et al. Genetic drift, selection and the evolution of the mutation rate. Nat Rev Genet. 2016; 17(11):704-14. doi: 10.1038/nrg.2016.104. PMID: 27739533.
23. Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A, et al. Tempo and mode of genome evolution in a 50,000-generation experiment. Nature. 2016; 536(7615):165-70. doi: 10.1038/nature18959. PMID: 27479321.
24. Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, et al. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature. 2013; 500(7464):571-4. doi: 10.1038/nature12344. PMID: 23873039.
25. Buskirk SW, Peace RE, Lang GI. Hitchhiking and epistasis give rise to cohort dynamics in adapting populations. Proc Natl Acad Sci U S A. 2017; 114(31):8330-5. doi: 10.1073/pnas.1702314114. PMID: 28720700.
26. Nguyen Ba AN, Cvijović I, Rojas Echenique JI, Lawrence KR, Rego-Costa A, Liu X, et al. High-resolution lineage tracking reveals travelling wave of adaptation in laboratory

yeast. Nature. 2019; 575(7783):494-9. doi: 10.1038/s41586-019-1749-3. PMID: 31723263.

27. McDonald MJ. Microbial Experimental Evolution - a proving ground for evolutionary theory and a tool for discovery. EMBO Rep. 2019; 20(8):e46992. doi: 10.15252/embr.201846992. PMID: 31338963.
28. Bailey SF, Alonso Morales LA, Kassen R. Effects of Synonymous mutations beyond codon bias: the evidence for adaptive synonymous substitutions from microbial evolution experiments. Genome Biol Evol. 2021; 13(9). doi: 10.1093/gbe/evab141. PMID: 34132772.
29. Agashe D. The evolutionary impacts of synonymous mutations. EcoEvoRxiv. 2021; https://doi.org/10.32942/osf.io/kn9cp.
30 Fronk DC, Sachs JL. Symbiotic organs: the nexus of host-microbe evolution. Trends Ecol Evol. 2022; 37(7):599-610. doi: 10.1016/j.tree.2022.02.014. PMID: 35393155.
31. Nyholm SV, McFall-Ngai MJ. The winnowing: establishing the squid-vibrio symbiosis. Nat Rev Microbiol. 2004; 2(8):632-42. doi: 10.1038/nrmicro957. PMID: 15263898.
32. Kumar N, Ganesh Lad, G, Giuntini E, Kaye M, Udomwong P, et al. Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum*. Open Biol. 2015; 5(1):140133. doi: 10.1098/rsob.140133. PMID: 25589577.
33. Daubech B, Remigi P, Doin de Moura G, Marchetti M, Pouzet C, Auriac MC, et al. Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation. Elife. 2017; 6. doi: 10.7554/eLife.28683. PMID: 29022875.
34. Joseph SB, Swanstrom R, Kashuba AD, Cohen MS. Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. Nat Rev Microbiol. 2015; 13(7):414-25. doi: 10.1038/nrmicro3471. PMID: 26052661;.
35. Hullahalli K, Waldor MK. Pathogen clonal expansion underlies multiorgan dissemination and organ-specific outcomes during murine systemic infection. Elife. 2021; 10. doi: 10.7554/eLife.70910. PMID: 34636322.
36. Goodrich-Blair H, Clarke DJ. Mutualism and pathogenesis in *Xenorhabdus* and *Photorhabdus*: two roads to the same destination. Mol Microbiol. 2007; 64(2):260-8. doi: 10.1111/j.1365-2958.2007.05671.x. PMID: 17493120.

37. Koinuma H, Maejima K, Tokuda R, Kitazawa Y, Nijo T, Wei W, et al. Spatiotemporal dynamics and quantitative analysis of phytoplasmas in insect vectors. Sci Rep. 2020; 10(1):4291. doi: 10.1038/s41598-020-61042-x. PMID: 32152370.
38. Perreau J, Zhang B, Maeda GP, Kirkpatrick M, Moran NA. Strong within-host selection in a maternally inherited obligate symbiont: *Buchnera* and aphids. Proc Natl Acad Sci U S A. 2021; 118(35). doi: 10.1073/pnas.2102467118. PMID: 34429360.
39. Bergstrom CT, McElhany P, Real LA. Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. Proc Natl Acad Sci U S A. 1999; 96(9):5095-100. doi: 10.1073/pnas.96.9.5095. PMID: 10220424.
40. LeClair J, Wahl L. The impact of population bottlenecks on microbial adaptation. J Stat Physics. 2018; 172:114–125, doi:10.1007/s10955-017-1924-6.
41. Izutsu M, Lake D, Matson Z, Dodson J, Lenski R. Effects of periodic bottlenecks on the dynamics of adaptive evolution in microbial populations. bioRxiv; 2022; doi:10.1101/2021.12.29.474457
42. Wein T, Dagan T. The Effect of population bottleneck size and selective regime on genetic diversity and evolvability in bacteria. Genome Biol Evol. 2019; 11(11):3283-90. doi: 10.1093/gbe/evz243. PMID: 31688900.
43. Garoff L, Pietsch F, Huseby DL, Lilja T, Brandis G, Hughes D. Population bottlenecks strongly influence the evolutionary trajectory to fluoroquinolone resistance in *Escherichia coli*. Mol Biol Evol. 2020; 37(6):1637-46. doi: 10.1093/molbev/msaa032. PMID: 32031639.
44. Mahrt N, Tietze A, Künzel S, Franzenburg S, Barbosa C, Jansen G, et al. Bottleneck size and selection level reproducibly impact evolution of antibiotic resistance. Nat Ecol Evol. 2021; 5(9):1233-42. doi: 10.1038/s41559-021-01511-2. PMID: 34312522.
45. Windels EM, Fox R, Yerramsetty K, Krouse K, Wenseleers T, Swinnen J, et al. Population bottlenecks strongly affect the evolutionary dynamics of antibiotic persistence. Mol Biol Evol. 2021; 38(8):3345-57. doi: 10.1093/molbev/msab107. PMID: 33871643.
46. Sousa A, Ramiro RS, Barroso-Batista J, Güleresi D, Lourenço M, Gordo I. Recurrent reverse evolution maintains polymorphism after strong bottlenecks in commensal gut bacteria. Mol Biol Evol. 2017; 34(11):2879-92. doi: 10.1093/molbev/msx221. PMID: 28961745.

47. Handel A, Bennett MR. Surviving the bottleneck: transmission mutants and the evolution of microbial populations. Genetics. 2008; 180(4):2193-200. doi: 10.1534/genetics.108.093013. PMID: 18854584.
48. Wilker PR, Dinis JM, Starrett G, Imai M, Hatta M, Nelson CW, et al. Selection on haemagglutinin imposes a bottleneck during mammalian transmission of reassortant H5N1 influenza viruses. Nat Commun. 2013; 4:2636. doi: 10.1038/ncomms3636. PMID: 24149915.
49. Moncla LH, Zhong G, Nelson CW, Dinis JM, Mutschler J, Hughes AL, et al. Selective bottlenecks shape evolutionary pathways taken during mammalian adaptation of a 1918-like avian *Influenza* virus. Cell Host Microbe. 2016; 19(2):169-80. doi: 10.1016/j.chom.2016.01.011. PMID: 26867176.
50. Sobel Leonard A, McClain MT, Smith GJ, Wentworth DE, Halpin RA, Lin X, et al. Deep sequencing of *Influenza* A virus from a human challenge study reveals a selective bottleneck and only limited intrahost genetic diversification. J Virol. 2016; 90(24):11247-58. doi: 10.1128/JVI.01657-16. PMID: 27707932.
51. Huang SW, Huang YH, Tsai HP, Kuo PH, Wang SM, Liu CC, et al. A selective bottleneck shapes the evolutionary mutant spectra of enterovirus A71 during viral dissemination in humans. J Virol. 2017; 91(23). doi: 10.1128/JVI.01062-17. PMID: 28931688.
52. Libby E, Rainey PB. Exclusion rules, bottlenecks and the evolution of stochastic phenotype switching. Proc Biol Sci. 2011; 278(1724):3574-83. doi: 10.1098/rspb.2011.0146. PMID: 21490013.
53. De Ste Croix M, Holmes J, Wanford JJ, Moxon ER, Oggioni MR, Bayliss CD. Selective and non-selective bottlenecks as drivers of the evolution of hypermutable bacterial loci. Mol Microbiol. 2020; 113(3):672-81. doi: 10.1111/mmi.14453. PMID: 32185830.
54. Schmutzer M, Wagner A. Gene expression noise can promote the fixation of beneficial mutations in fluctuating environments. PLoS Comput Biol. 2020; 16(10):e1007727. doi: 10.1371/journal.pcbi.1007727. PMID: 33104710.
55. Gama JA, Abby SS, Vieira-Silva S, Dionisio F, Rocha EP. Immune subversion and quorum-sensing shape the variation in infectious dose among bacterial pathogens. PLoS Pathog. 2012; 8(2):e1002503. doi: 10.1371/journal.ppat.1002503. PMID: 22319444.

56. Abel S, zur Wiesch, PA, Davis B, Waldor M. Analysis of bottlenecks in experimental models of infection. PLoS Pathogen; 2015; 11(6):e1004823doi:10.1371/journal.ppat.1004823. PMID: 26066486.

57. Papkou A, Gokhale CS, Traulsen A, Schulenburg, H. Host–parasite coevolution: why changing population size matters. Zoology; 2016; 119(4):330-8. doi: 10.1016/j.zool.2016.02.001. PMID: 27161157.

58. Cossart P, Sansonetti PJ. Bacterial invasion: the paradigms of enteroinvasive pathogens. Science. 2004; 304(5668):242-8. doi: 10.1126/science.1090124. PMID: 15073367.

59. Visick KL, Stabb EV, Ruby EG. A lasting symbiosis: how *Vibrio fischeri* finds a squid partner and persists within its natural host. Nat Rev Microbiol. 2021; 19(10):654-65. doi: 10.1038/s41579-021-00557-0. PMID: 34089008.

60. Rowe HM, Karlsson E, Echlin H, Chang TC, Wang L, van Opijnen T, et al. Bacterial Factors Required for Transmission of Streptococcus pneumoniae in Mammalian Hosts. Cell Host Microbe. 2019; 25(6):884-91.e6. doi: 10.1016/j.chom.2019.04.012. PMID: 31126758.

61. Bright M, Bulgheresi S. A complex journey: transmission of microbial symbionts. Nat Rev Microbiol. 2010; 8(3):218-30. doi: 10.1038/nrmicro2262. PMID: 20157340.

62. Bongrand C, Moriano-Gutierrez S, Arevalo P, McFall-Ngai M, Visick KL, Polz M, et al. Using colonization assays and comparative genomics to discover symbiosis behaviors and factors in *Vibrio fischeri*. mBio. 2020; 11(2). doi: 10.1128/mBio.03407-19. PMID: 32127462.

63. Mendoza-Suárez M, Andersen SU, Poole PS, Sánchez-Cañizares C. Competition, nodule occupancy, and persistence of inoculant strains: key factors in the rhizobium-legume symbioses. Front Plant Sci. 2021; 12:690567. doi: 10.3389/fpls.2021.690567. PMID: 34489993.

64. Robinson CD, Klein HS, Murphy KD, Parthasarathy R, Guillemin K, Bohannan BJM. Experimental bacterial adaptation to the zebrafish gut reveals a primary role for immigration. PLoS Biol. 2018; 16(12):e2006893. doi: 10.1371/journal.pbio.2006893. PMID: 30532251.

65. Bacigalupe R, Tormo-Mas M, Penadés JR, Fitzgerald JR. A multihost bacterial pathogen overcomes continuous population bottlenecks to adapt to new host species. Sci Adv. 2019; 5(11):eaax0063. doi: 10.1126/sciadv.aax0063. PMID: 31807698.

66. Ebert D, Duneau D, Hall MD, Luijckx P, Andras JP, Du Pasquier L, et al. A population biology perspective on the stepwise infection process of the bacterial pathogen *Pasteuria ramosa* in *Daphnia*. Adv Parasitol. 2016; 91:265-310. doi: 10.1016/bs.apar.2015.10.001. PMID: 27015951.
67. Hall MD, Routtu J, Ebert D. Dissecting the genetic architecture of a stepwise infection process. Mol Ecol. 2019; 28(17):3942-57. doi: 10.1111/mec.15166. PMID: 31283079.
68. Xiao TT, Schilderink S, Moling S, Deinum EE, Kondorosi E, Franssen H, et al. Fate map of *Medicago truncatula* root nodules. Development. 2014; 141(18):3517-28. doi: 10.1242/dev.110775. PMID: 25183870.
69. Moling S, Pietraszewska-Bogiel A, Postma M, Fedorova E, Hink MA, Limpens E, et al. Nod factor receptors form heteromeric complexes and are essential for intracellular infection in *Medicago* nodules. Plant Cell. 2014; 26(10):4188-99. doi: 10.1105/tpc.114.129502. PMID: 25351493.
70. Cerri MR, Frances L, Kelner A, Fournier J, Middleton PH, Auriac MC, et al. The symbiosis-related ERN transcription factors act in concert to coordinate rhizobial host root infection. Plant Physiol. 2016; 171(2):1037-54. doi: 10.1104/pp.16.00230. PMID: 27208242.
71. Capoen W, Goormachtig S, De Rycke R, Schroeyers K, Holsters M. SrSymRK, a plant receptor essential for symbiosome formation. Proc Natl Acad Sci U S A. 2005; 102(29):10369-74. doi: 10.1073/pnas.0504250102. PMID: 16006516.
72. Liu J, Rasing M, Zeng T, Klein J, Kulikova O, Bisseling T. NIN is essential for development of symbiosomes, suppression of defence and premature senescence in *Medicago truncatula* nodules. New Phytol. 2021; 230(1):290-303. doi: 10.1111/nph.17215. PMID: 33471433.
73. Bovin S, Lepetit M. Partner preference in the legume-rhizobia symbiosis and impact on legume inoculation strategies. Frendo P, Frugier F, Masson-Boivin C ed: Advances in Botanical Research. 2020; 94, 323-348, doi:10.1016/bs.abr.2019.09.016.
74. Younginger BS, Friesen ML. Connecting signals and benefits through partner choice in plant-microbe interactions. FEMS Microbiol Lett. 2019; 366(18). doi: 10.1093/femsle/fnz217. PMID: 31730203.

75. Batstone RT, O'Brien AM, Harrison TL, Frederickson ME. Experimental evolution makes microbes more cooperative with their local host genotype. Science. 2020; 370(6515):476-8. doi: 10.1126/science.abb7222. PMID: 33093112.

76. Quides KW, Weisberg AJ, Trinh J, Salaheldine F, Cardenas P, Lee HH, et al. Experimental evolution can enhance benefits of rhizobia to novel legume hosts. Proc Biol Sci. 2021; 288(1951):20210812. doi: 10.1098/rspb.2021.0812. PMID: 34034525.

77. Pankey SM, Foxall RL, Ster IM, Perry LA, Schuster BM, Donner RA, et al. Host-selected mutations converging on a global regulator drive an adaptive leap towards symbiosis in bacteria. Elife. 2017; 6. doi: 10.7554/eLife.24414. PMID: 28447935.

78. Faucher C, Mazana V, Kardacz M, Parthuisot N, Ferdy JB, Duneau D. Step-specific adaptation and trade-off over the course of an infection by GASP mutation small colony variants. mBio. 2021; 12(1). doi: 10.1128/mBio.01399-20. PMID: 33436427.

79. Cambon MC, Parthuisot N, Pagès S, Lanois A, Givaudan A, Ferdy JB. Selection of bacterial mutants in late infections: when vector transmission trades off against growth advantage in stationary phase. mBio. 2019; 10(5). doi: 10.1128/mBio.01437-19. PMID: 31594811.

80. Costa G, Gildenhard M, Eldering M, Lindquist RL, Hauser AE, Sauerwein R, et al. Non-competitive resource exploitation within mosquito shapes within-host malaria infectivity and virulence. Nat Commun. 2018; 9(1):3474. doi: 10.1038/s41467-018-05893-z. PMID: 30150763.

81. Van den Bergh B, Swings T, Fauvart M, Michiels J. Experimental design, population dynamics, and diversity in microbial experimental evolution. Microbiol Mol Biol Rev. 2018; 82(3). doi: 10.1128/MMBR.00008-18. PMID: 30045954.

82. Couce A, Tenaillon OA. The rule of declining adaptability in microbial evolution experiments. Front Genet. 2015;6:99. doi: 10.3389/fgene.2015.00099. PMID: 25815007.

83. Maddamsetti R, Lenski RE, Barrick JE. Adaptation, clonal interference, and frequency-dependent interactions in a long-term evolution experiment with *Escherichia coli*. Genetics. 2015; 200(2):619-31. doi: 10.1534/genetics.115.176677. PMID: 25911659.

84. Neher R. Genetic draft, selective interference, and population genetics of rapid adaptation. Ann Rev Ecol Evol Syst. 2013; 44:195-215. doi:10.1146/annurev-ecolsys-110512-135920.

85. Katsnelson MI, Wolf YI, Koonin EV. On the feasibility of saltational evolution. Proc Natl Acad Sci U S A. 2019; 116(42):21068-75. doi: 10.1073/pnas.1909031116. PMID: 31570621.

86. Beringer. R-factor transfer in *Rhizobium leguminosarum*. J Gen Microbiol. 1974; 84(1):188-98. doi: 10.1099/00221287-84-1-188. PMID: 4612098.

87. Boucher CA, Barberis PA, Trigalet AP, Demery DA. Transposon mutagenesis of *Pseudomonas solanacearum* isolation of Tn5-induced avirulent mutants. J Gen Microbiol. 1985; 131(SEP):2449-57.

88. 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(4):1011-9. doi: 10.1128/jb.01189-09. PMID: 20008073.

89. Bertani G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J Bacteriol. 1951; 62(3):293-300. doi: 10.1128/jb.62.3.293-300.1951. PMID: 14888646.

90. Deatherage DE, Barrick JE. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods Mol Biol. 2014; 1151:165-88. doi: 10.1007/978-1-4939-0554-6_12. PMID: 24838886.

91. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011; 29(1):24-6. doi: 10.1038/nbt.1754. PMID: 21221095.

92. Vallenet D, Calteau A, Dubois M, Amours P, Bazin A, Beuvin M, et al. MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. Nucleic Acids Res. 2020; 48(D1):D579-D89. doi: 10.1093/nar/gkz926. PMID: 31647104.

93. Team RDC. R: a language and environment for statistical computing. 2008.

94. Noble R. ggmuller: Create Muller Plots of Evolutionary Dynamics. R package version 0.5.3. 2019.

95. Dalia AB, McDonough E, Camilli A. Multiplex genome editing by natural transformation. Proc Natl Acad Sci U S A. 2014; 111(24):8937-42. doi: 10.1073/pnas.1406478111. PMID: 24889608.

96. Chen WM, Moulin L, Bontemps C, Vandamme P, Bena G, Boivin-Masson C. Legume symbiotic nitrogen fixation by beta-proteobacteria is widespread in nature. J Bacteriol. 2003; 185(24):7266-72. doi: 10.1128/jb.185.24.7266-7272.2003. PMID: 14645288.
97. Fahraeus G. The infection of Clover root hairs by nodule bacteria studied by a simple glass slide technique. J Gen Microbiol. 1957; 16(2):374-381. PMID: 13416514.
98. Jensen. Nitrogen fixation in leguminous plants. I. General characters of root nodule bacteria isolated from species of *Medicago* and *Trifolium* in Australia. Proc Int Soc N S W. 1942; p. 68–108.
99. Doin de Moura GG, Remigi P, Masson-Boivin C, Capela D. Experimental evolution of legume symbionts: what have we learnt? Genes (Basel). 2020; 11(3). doi: 10.3390/genes11030339. PMID: 32210028.
100. Tenaillon O, Toupance B, Le Nagard H, Taddei F, Godelle B. Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. Genetics. 1999; 152(2):485-93. PMID: 10353893.
101. Tenaillon O. The Utility of Fisher's Geometric Model in Evolutionary Genetics. Annu Rev Ecol Evol Syst. 2014; 45:179-201. doi: 10.1146/annurev-ecolsys-120213-091846. PMID: 26740803.
102. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019; https://www.R-project.org/.
103. Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F. and Hothorn, T. mvtnorm: multivariate normal and t distributions. R package version 1.1-1. 2020; http://CRAN.Rproject.org/package=mvtnorm.
104. Ren, K. rlist: A toolbox for non-tabular data manipulation. R package version 0.4.6.1. 2016; https://CRAN.R-project.org/package=rlist.
105. Wickham, H. ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. 2016; https://ggplot2.tidyverse.org.
106. Wickham, H. et al., Welcome to the tidyverse. J Open Source Software. 2019; 4:1686. https://doi.org/10.21105/joss.01686.
107. Wickham, H. stringr: Simple, Consistent wrappers for common string operations. R package version 1.4.0. 2019; https://CRAN.R-project.org/package=stringr.

108. Alizon S, Michalakis Y. Adaptive virulence evolution: the good old fitness-based approach. Trends Ecol Evol. 2015; 30(5):248-54. doi: 10.1016/j.tree.2015.02.009. PMID: 25837917.

Supporting information

S1 Fig. Experimental evolution of *Ralstonia solanacearum* GMI1000 pRalta through serial cycles of plant (*Mimosa pudica*) inoculation-isolation of nodule bacteria.

S2 Fig. Dry weights of *M. pudica* plants inoculated with cycle 35 evolved clones. The data underlying S2 Fig can be found in S13 Table

S3 Fig. Evolution of population composition over time in lines B and G.

S4 Fig. Survival of reconstructed mutants from lineages B and G in the Jensen medium and rhizosphere. The data underlying S4 Fig can be found in S17 Table.

S5 Fig. Schematic representation of the modelling framework.

S6 Fig. Distribution of fitness effects of new mutations for each of the two fitness components (nodulation competitiveness and proliferation), computed for three phenotypic dimensions.

S7 Fig. Representative distributions of fitness effects of new mutations for three levels of pleiotropy between nodulation competitiveness and within-host proliferation.

S8 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation (evolutionary parameters: k = l = 10 and m = 0).

S9 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation, with a higher probability of beneficial mutations (evolutionary parameters: k = l = 3 and m = 0).

S10 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation, with a lower probability of beneficial mutations (evolutionary parameters: k = l = 20 and m = 0).

S11 Fig. Effect of the fitness of the ancestor on the relative strength of selection for nodulation competitiveness and proliferation (evolutionary parameters: k = l = 10 and m = 0).

S12 Fig. Effect of the fitness of the ancestor on the relative strength of selection for nodulation competitiveness and proliferation, with a higher probability of beneficial mutations (evolutionary parameters: k = l = 3 and m = 0).

S13 Fig. Effect of the fitness of the ancestor on the relative strength of selection for nodulation competitiveness and proliferation, with a lower probability of beneficial mutations (evolutionary parameters: k = l = 20 and m = 0).

S14 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation under weak partial pleiotropy (evolutionary parameters: k = l = 6 and m = 4).

S15 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation under strong partial pleiotropy (evolutionary parameters: k = l = 2 and m = 8).

S16 Fig.: Effect of the chronology of symbiotic events and the size of nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation (evolutionary parameters: k = l = 10 and m = 0).

S1 Table. Number of nodules harvested at each evolution cycle

S2 Table. Summary of detected mutations in the five lineages

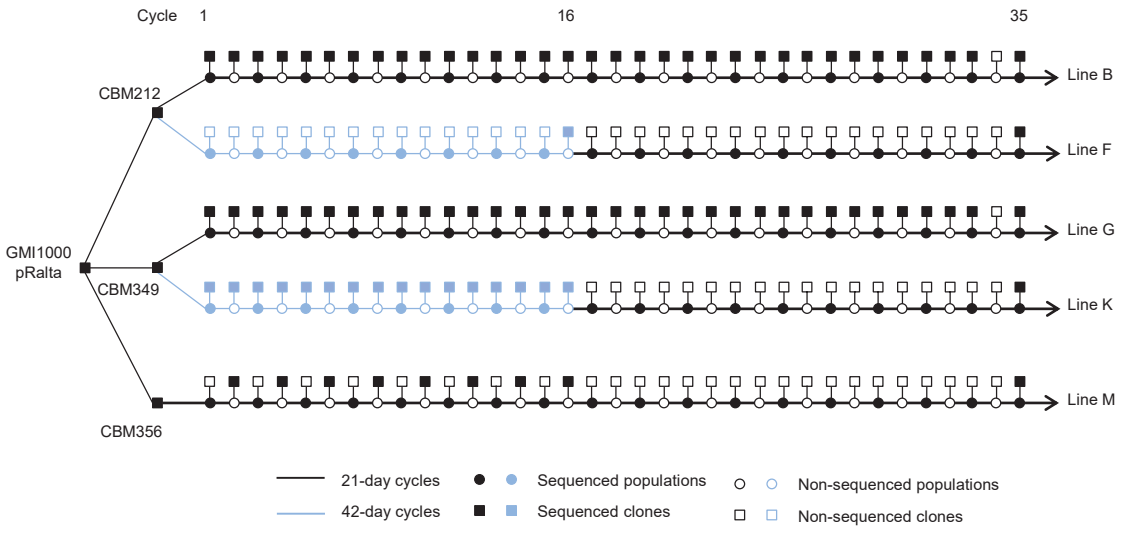
S3 Table. Global analysis of mutational cohorts

S4 Table. Neutral and deleterious reconstructed mutations

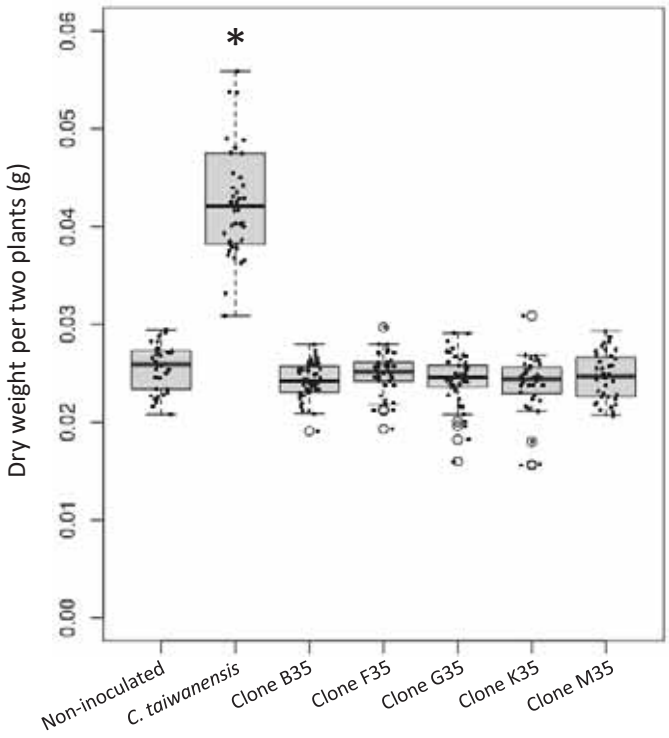
S5 Table. Frequencies of usage of codons modified by adaptive synonymous mutations

S6 Table. Strains, plasmids, oligonucleotides and softwares used in this study

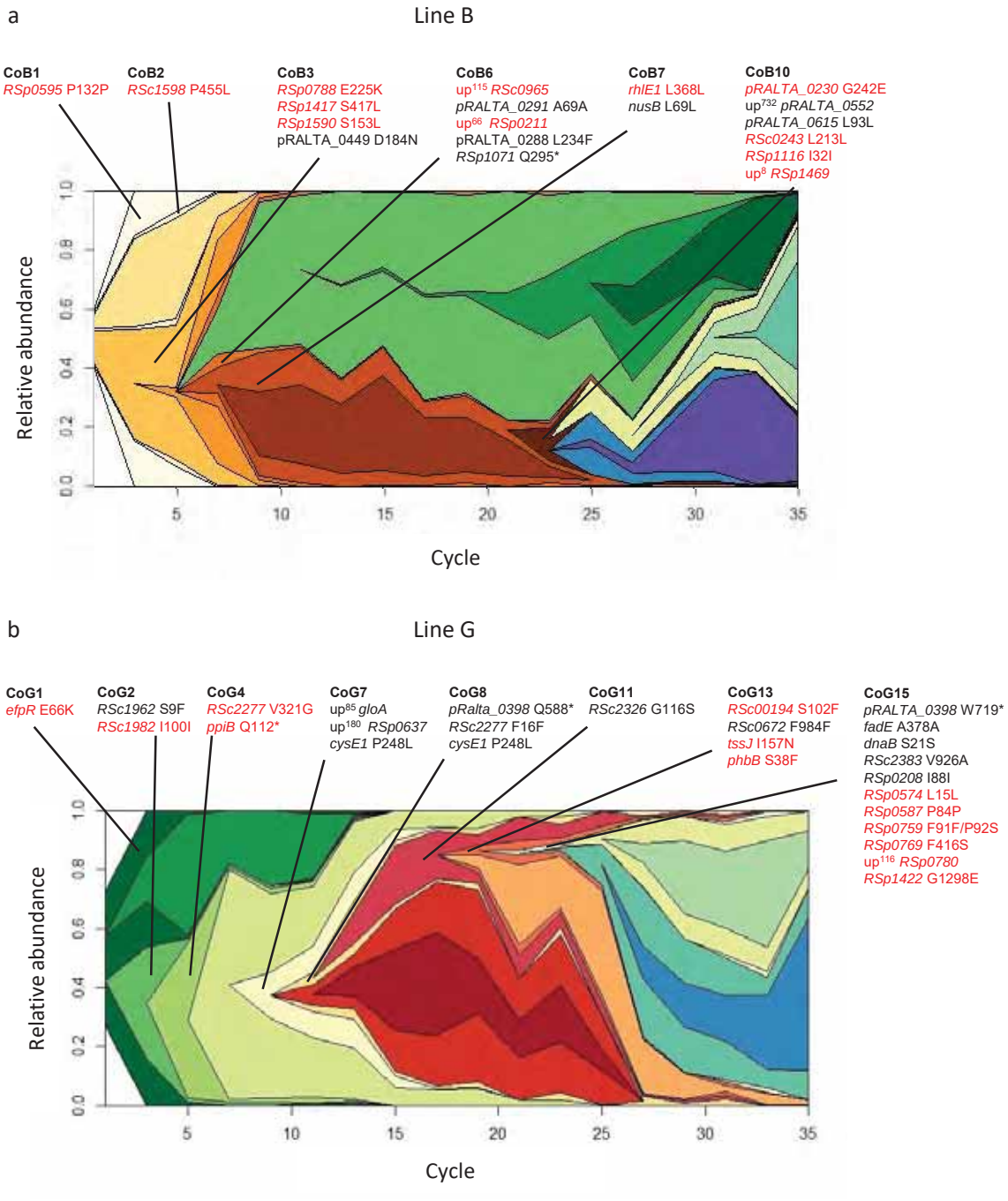
S7 Table. Parameter values used in computer simulations



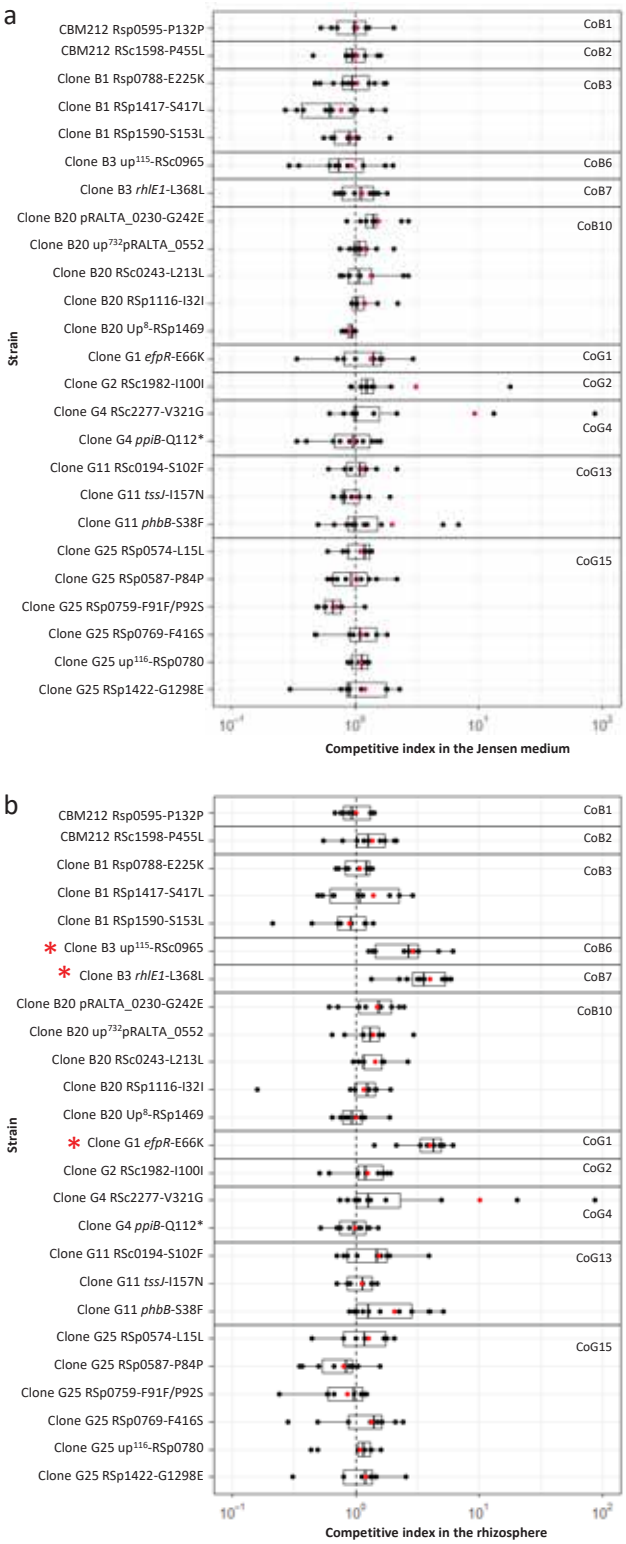
S2.1 Fig. Experimental evolution of *Ralstonia solanacearum* GMI1000 pRalta through serial cycles of plant (*Mimosa pudica*) inoculation-isolation of nodule bacteria.
The 5 bacterial lineages, lines B, F, G, K, and M were independently derived from the spontaneous nodulating mutants either CBM212, or CBM349, or CBM356. At each evolution cycle, bacterial nodule populations were isolated either 21 days (black lines) or 42 days (blue lines) after plant inoculation and partly directly re-inoculated to new plants to initiate the next cycle. At each cycle, one individual clone representative of the bacterial population was isolated and stored. Sequenced populations and clones are indicated.



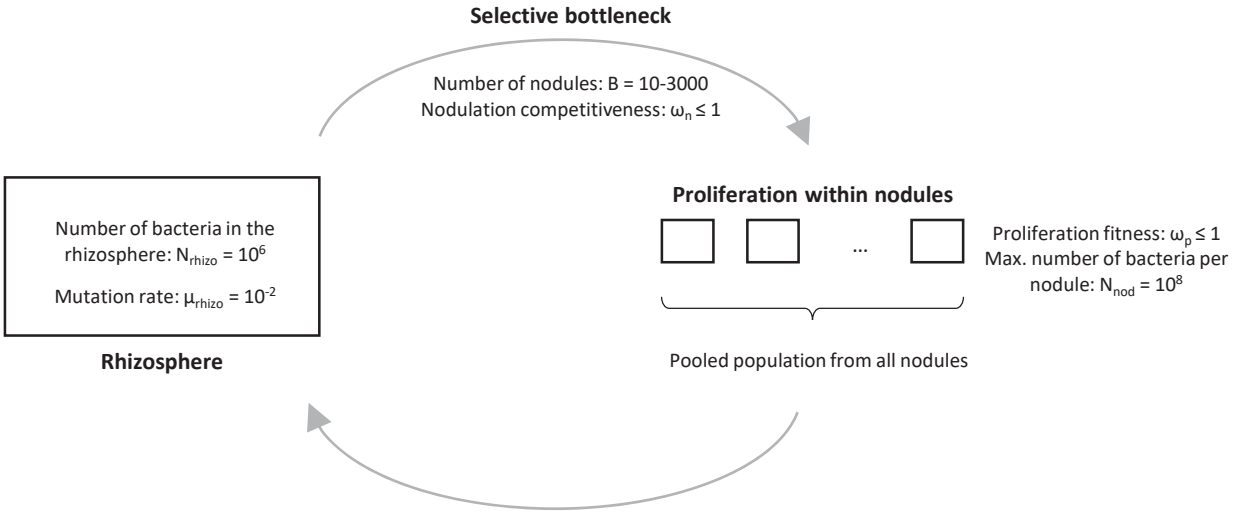
S2.2 Fig. Dry weights of *M. pudica* plants inoculated with cycle 35 evolved clones.
The dry weight of the aerial parts of two plants was measured 21 days after inoculation with the evolved clones of cycle 35 or *C. taiwanensis* and compared to non inoculated plants. Data were obtained from 3 to 4 independent experiments for each comparison and the sample size (n) is comprised between n=38-49 . * Significantly different from non inoculated plants ($P<0.05$, Wilcoxon test).



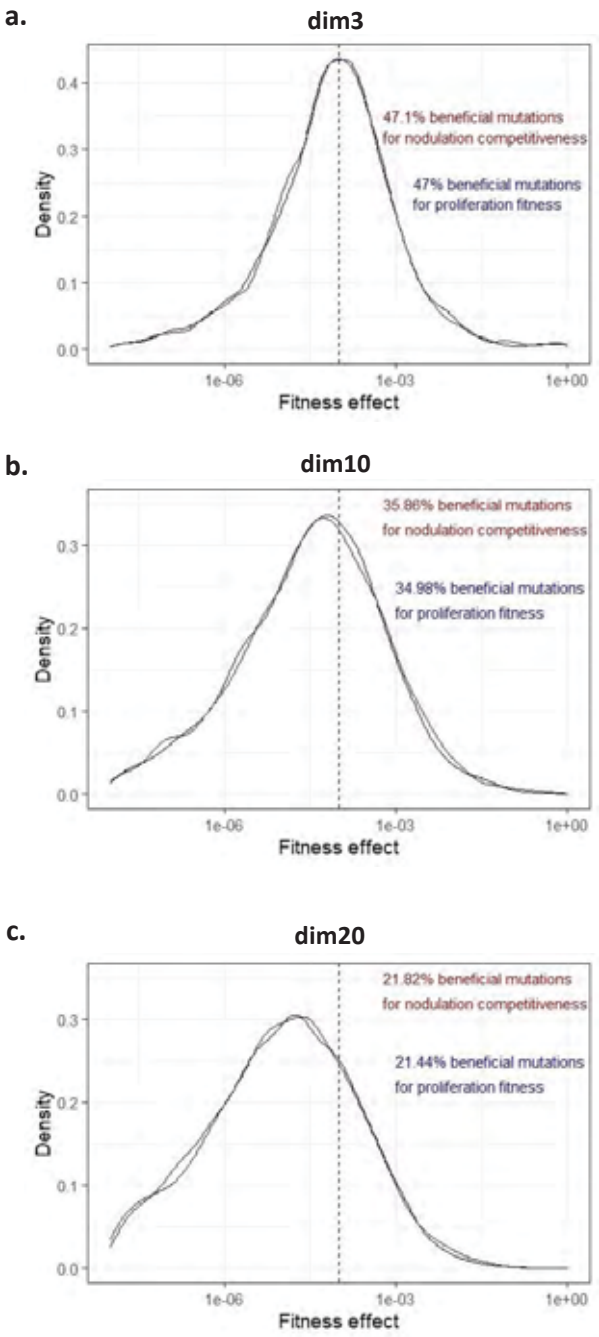
S2.3 Fig. Evolution of population composition over time in lines B and G.
Muller plots representing the relative frequency of each genotype in lines B (a) and G (b) for each sequenced population (uneven cycles). Each color represent a different genotype, and its relative frequency is shown as the vertical area at each sequenced time point. The genotypes containing cohorts for which adaptive mutations were identified experimentally are indicated. For each cohort, non-adaptive or deleterious mutations are shown in black, adaptive mutations are in red.



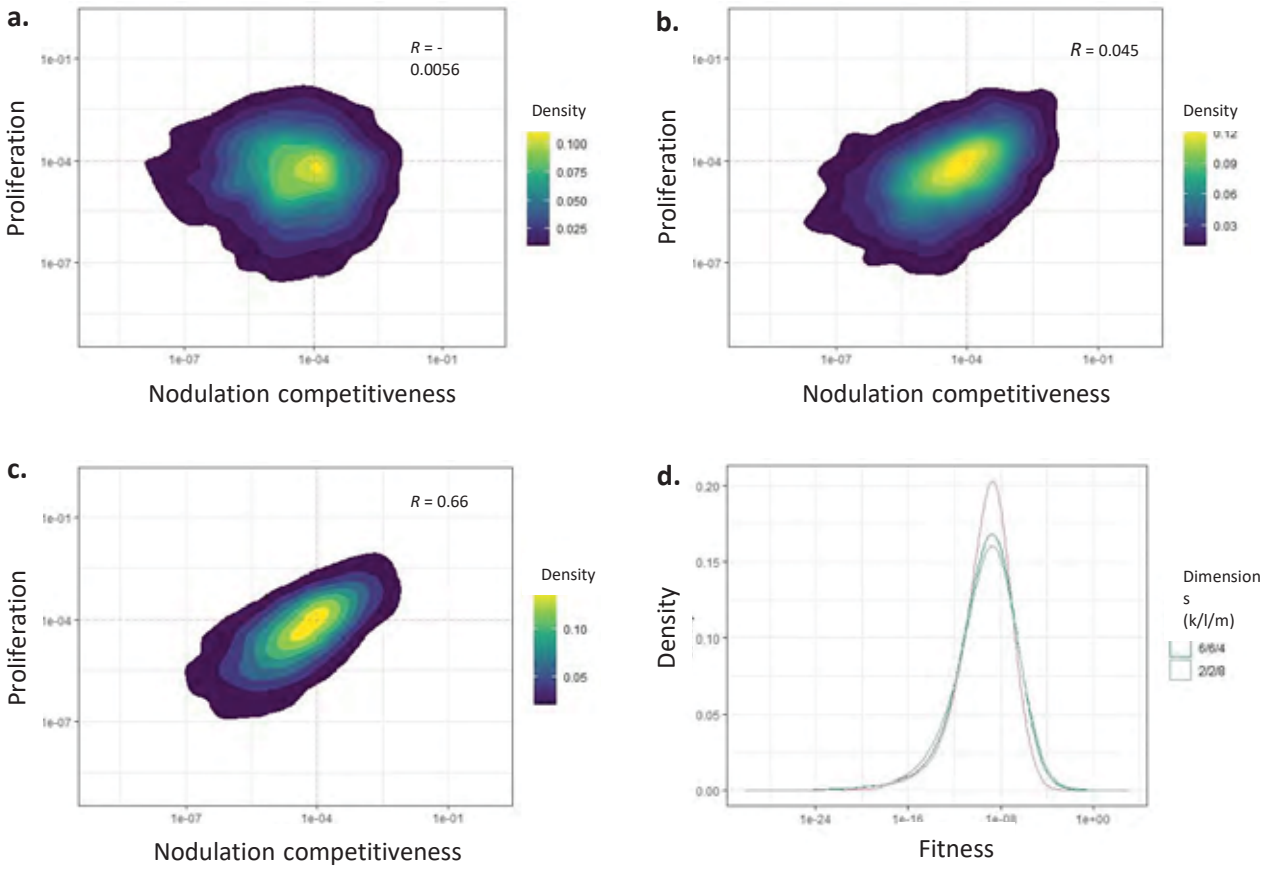
S2.4 Fig. Survival of reconstructed mutants from lineages B and G in the Jensen medium and rhizosphere.
Survival of evolved clones carrying reconstructed adaptive mutations from fixed mutational cohorts in the Jensen medium (**a**) and in rhizosphere (**b**). Competitive indexes (CI) were calculated as the ratio of the mutant strain on the isogenic parental strain in Jensen medium populations normalized by the inoculum ratio. Red points correspond to mean values of CI. Rectangles span the first quartiles to the third quartiles, bold segments inside the rectangle show the median. Red stars beside strain names indicate the mutations that significantly improved survival in the rhizosphere compared to the parental evolved clone ($P<0.05$, Wilcoxon test with Benjamini-Hochberg correction). Data were obtained from 3-5 independent experiments for each comparison and the sample size (n) is comprised between n=9-14.



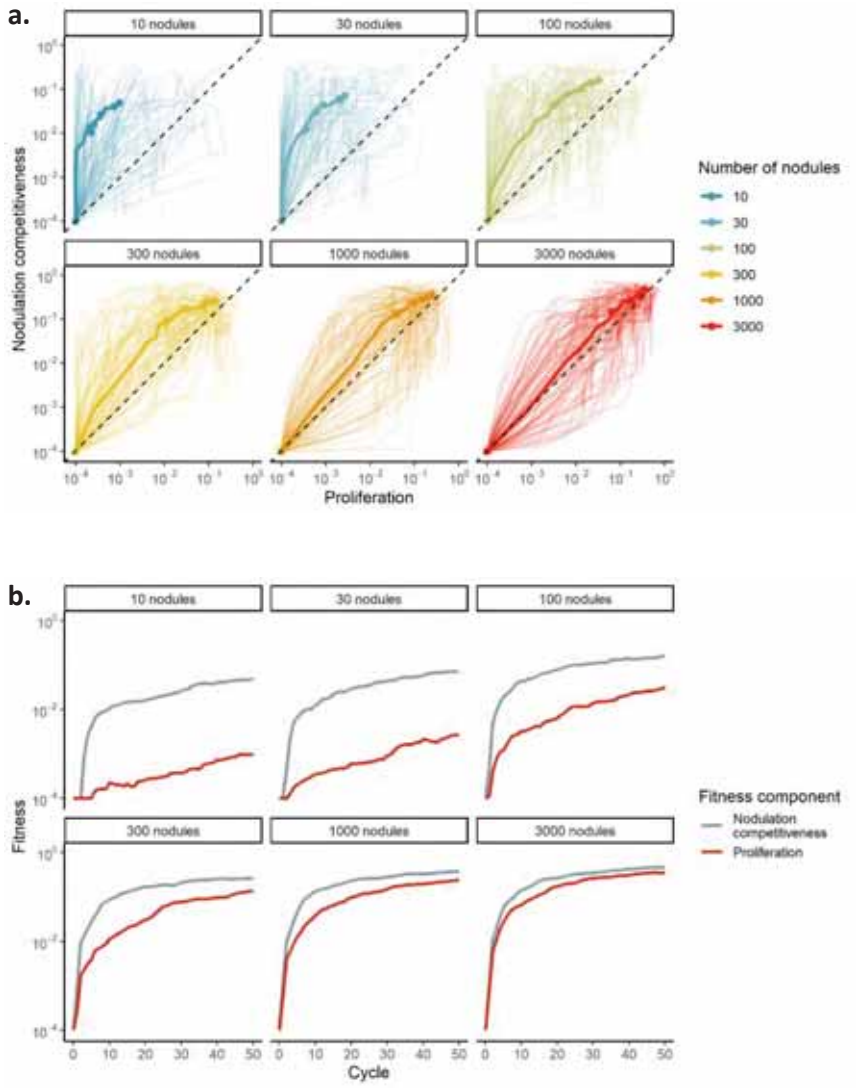
S2.5 Fig. Schematic representation of the modelling framework.
In the rhizosphere, bacteria acquire mutations (at a rate μ_{rhizo}) that will modify the value of the two fitness components: nodulation competitiveness ω_n) and proliferation capacity within nodules (ω_p). To enter the host, bacteria go through a selective bottleneck, where the probability that each bacterium founds one nodule is equal to its relative nodulation competitiveness value. At each cycle, a fixed number of nodules (B) are formed, defining the stringency of the bottleneck. Once within nodules, bacteria proliferate to a level equal to the product of their proliferation fitness value (ω_p) times the maximal number of bacteria per nodule (N_{nod}). At the end of each cycle, all bacteria from all nodules are pooled, and a subset of this population is used to found the rhizospheric population for the next cycle. This step creates a weak non-selective bottleneck with no effect on the evolutionary dynamics since all nodulating clones are represented according to their relative frequency in the nodule population ($B << N_{rhizo}$).



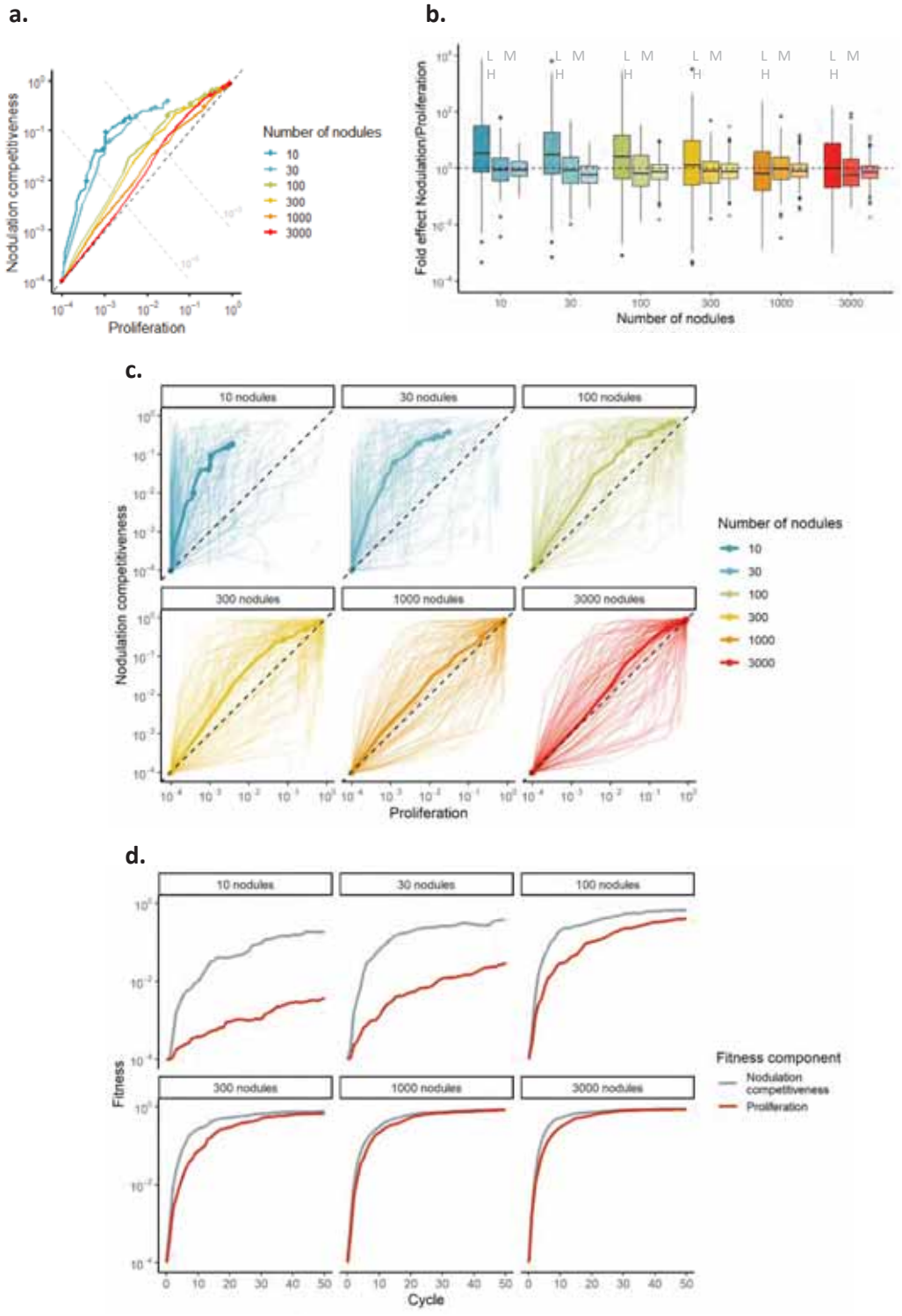
S2.6 Fig. Distribution of fitness effects of new mutations for each of the two fitness components (nodulation competitivity and proliferation), computed for three phenotypic dimensions: k = l = 3 (**a**), k = l = 10 (**b**) or k = l = 20 (**c**). In all cases, there is no pleiotropy (m = 0). Distributions were estimated by simulating 1 random mutation in each of 5000 different genetic architectures in the ancestral strain with low fitness (10^{-4} relative to the theoretical optimum for each fitness component, indicated by the vertical dashed line). On each panel, the percentage of beneficial mutations for each component is shown. Differences in the percentage of beneficial mutations between the two fitness components result from the randomness in the definition of the variance-covariance matrices of mutation and fitness effects.



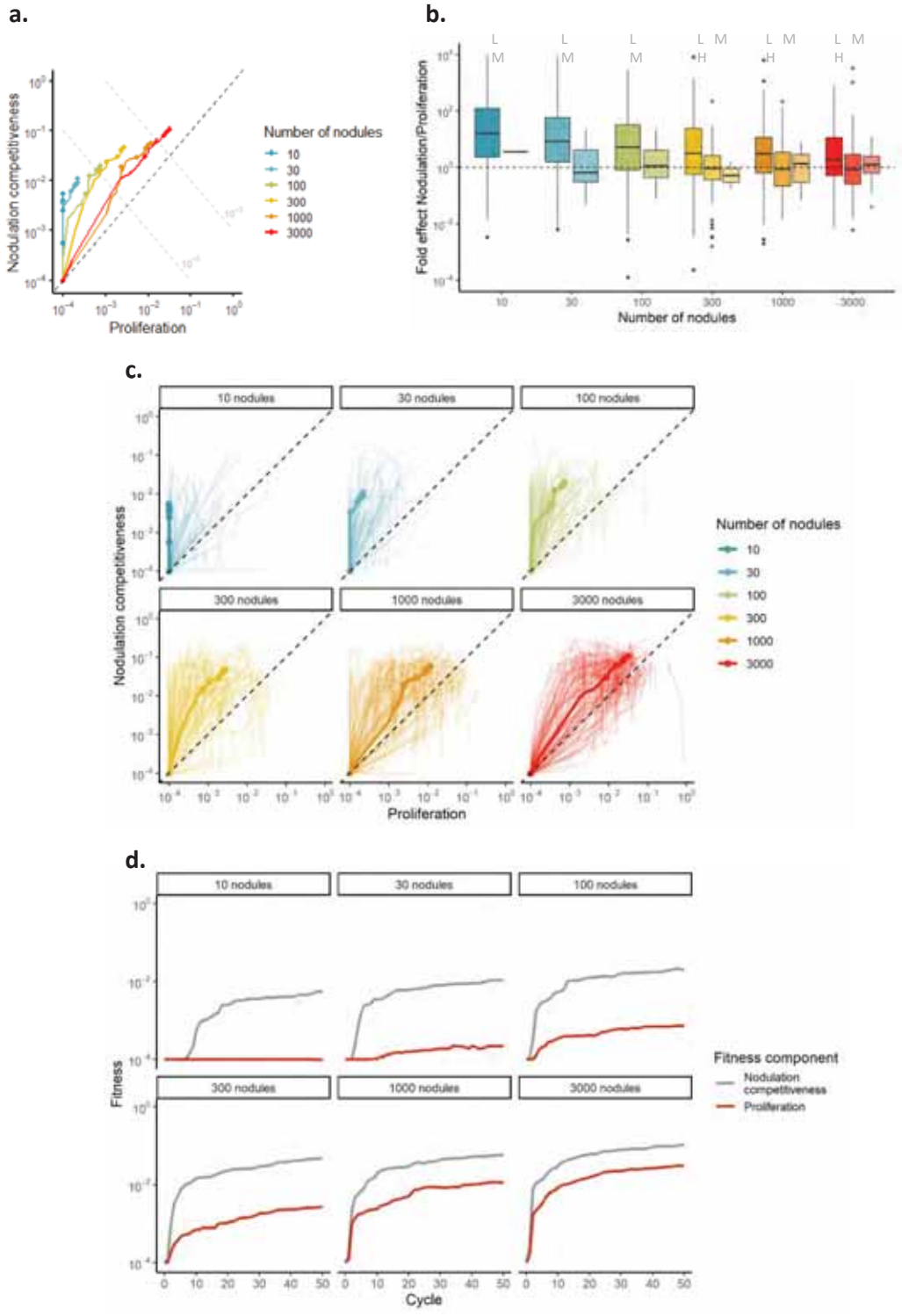
S2.7 Fig. Representative distributions of fitness effects of new mutations for three levels of pleiotropy between nodulation competitiveness and within-host proliferation. **a**, No pleiotropy (k = 10, l = 10, m = 0). **b**, Weak partial pleiotropy (k = 6, l = 6, m = 4). **c**, Stronger partial pleiotropy (k = l = 2, m = 8). **d**, Distribution of fitness effects (calculated as the product of nodulation competitiveness by proliferation) for simulated mutations shown in **a-c.** Distributions were computed by simulating 1 mutation in each of 5000 different genetic architectures in the ancestral strain with low fitness (10^{-4} relative to the theoretical optimum for each fitness component, indicated by the red dotted lines).



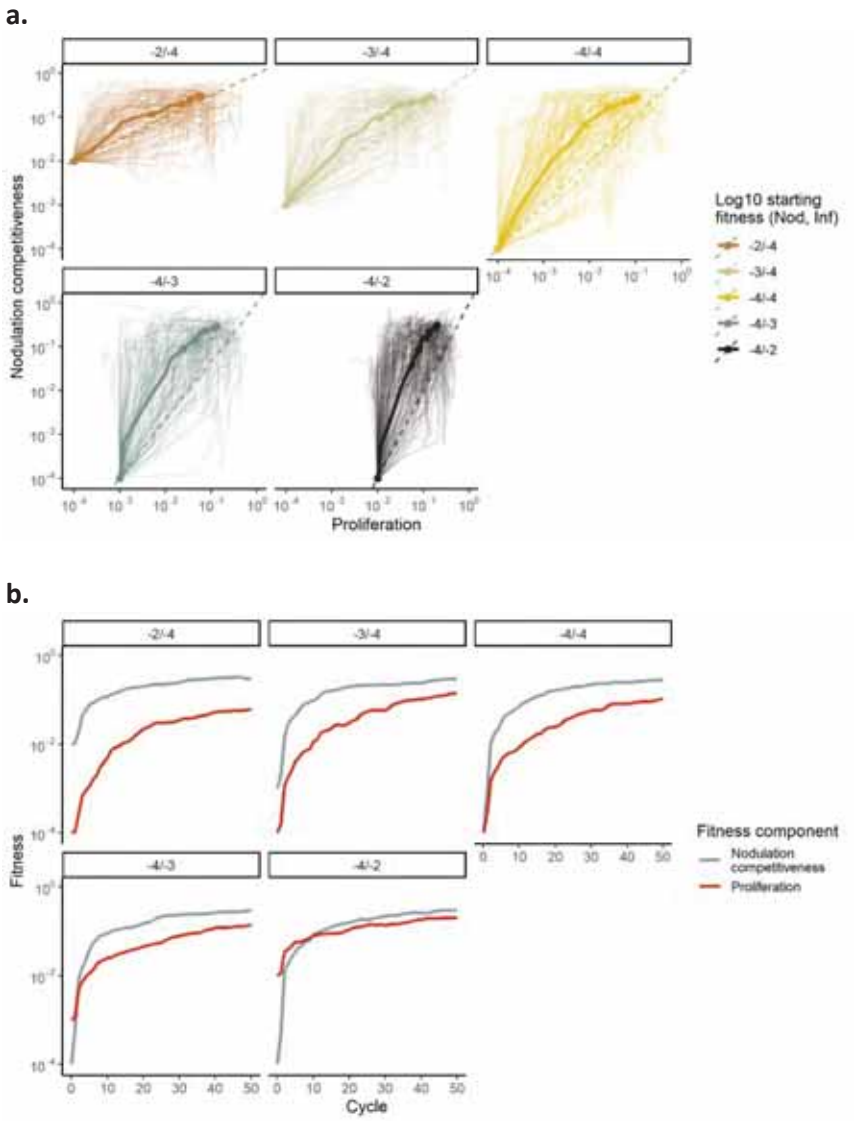
S2.8 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation. Same data as in Fig 6a,b (main text). **a**, Median (thick lines) and individual (thin lines) fitness trajectories of 100 simulated populations evolving under different sizes of nodulation bottleneck (10 to 3000 nodules). The black dotted line represents the diagonal, along which population would improve both phenotypic traits equally well. **b**, Improvement of nodulation competitiveness (grey lines) and proliferation (red lines) fitness values along evolutionary cycles under different sizes of nodulation bottlenecks. Evolutionary parameters used were: k = l = 10 and m = 0.



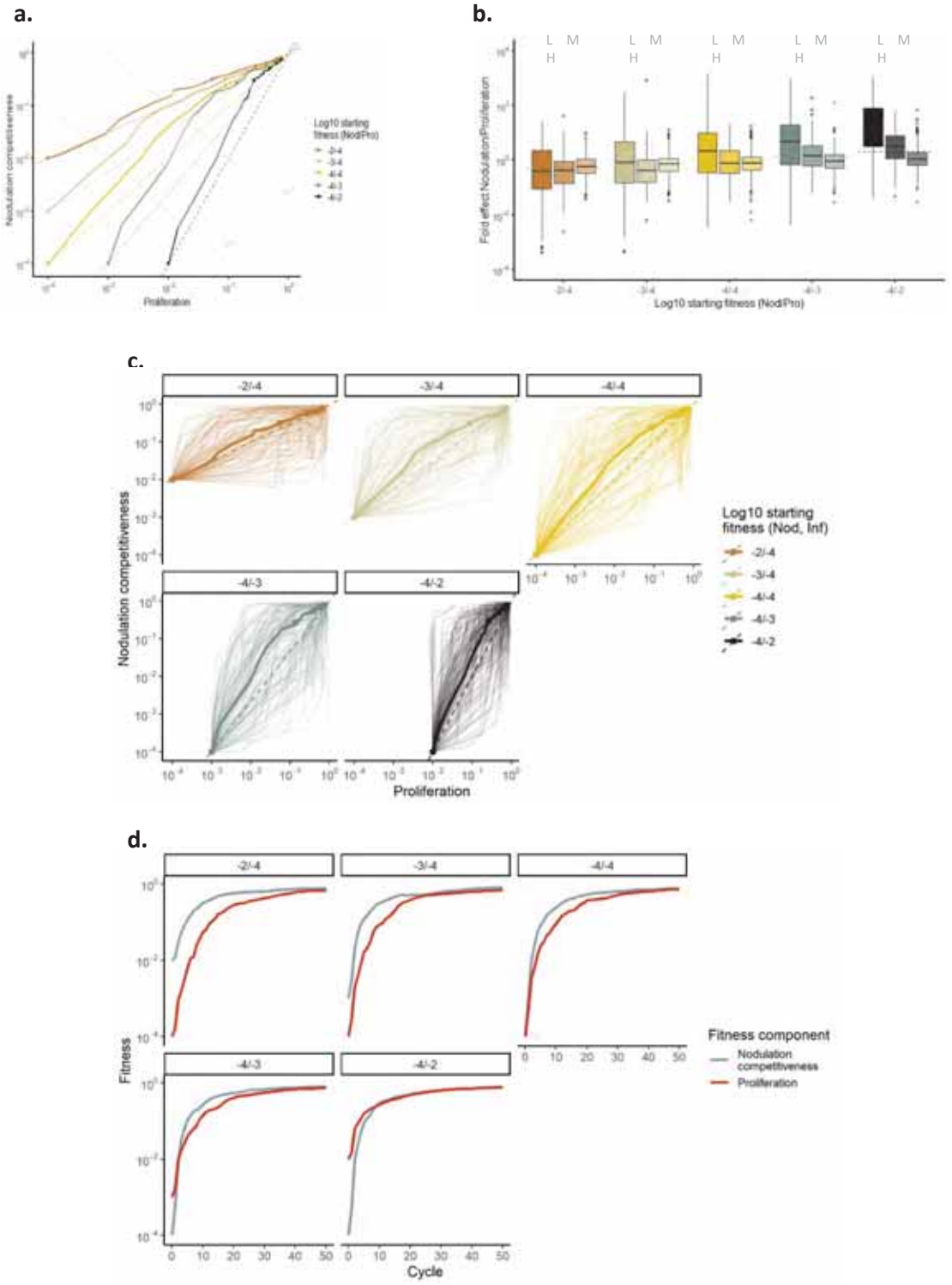
S2.9 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation, with a higher probability of beneficial mutations. **a,b**, as in Fig 6a,b (main text). **c,d,** as in S8 Fig. Evolutionary parameters used were: k = l = 3 and m = 0.



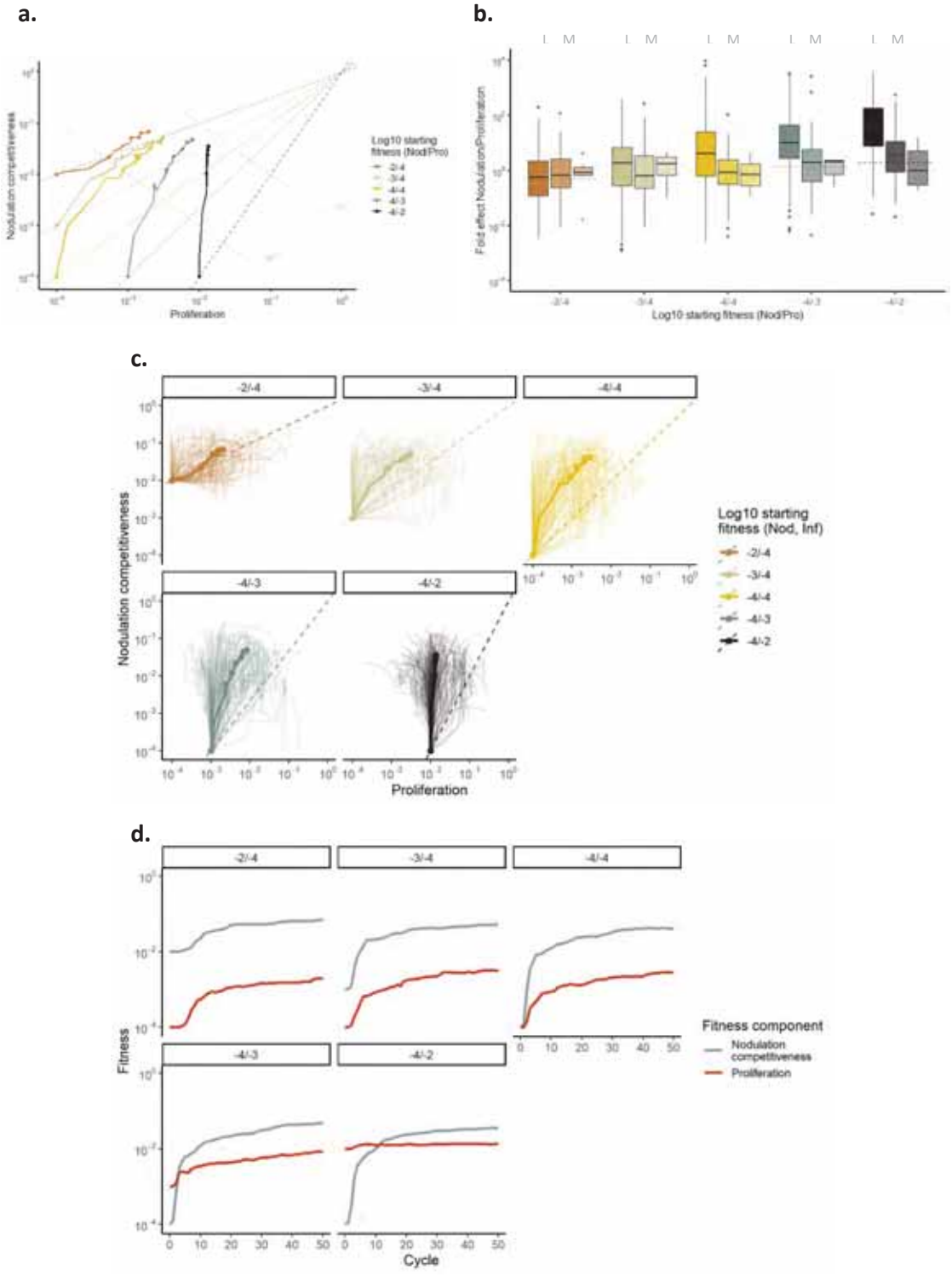
S2.10 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation, with a lower probability of beneficial mutations. **a,b,** as in Fig 6a,b (main text). **c,d**, as in S8 Fig. Evolutionary parameters used were: k = l = 20 and m = 0.



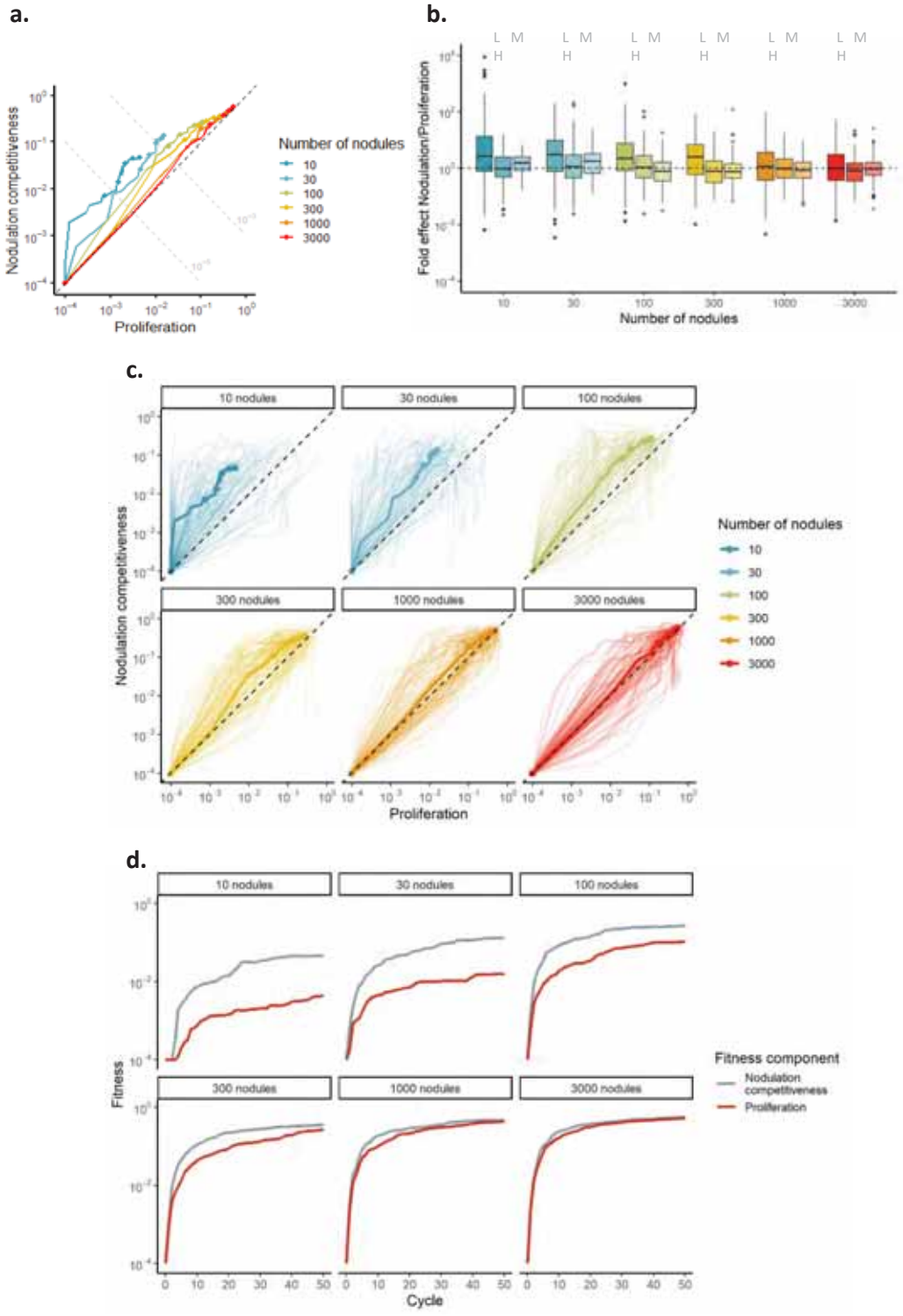
S2.11 Fig. Effect of the fitness of the ancestor on the relative strength of selection for nodulation competitiveness and proliferation. Same data as in Fig 6c,d. **a,b,** as in S8 Fig. , with fitness trajectories shown for various combinations of nodulation competitiveness/proliferations fitness values in the ancestor ($10^{-2}/10^{-4}$), ($10^{-2}/10^{-4}$), ($10^{-4}/10^{-4}$), ($10^{-4}/10^{-3}$), and ($10^{-4}/10^{-2}$). Evolutionary parameters used were: k = l = 10 and m = 0.



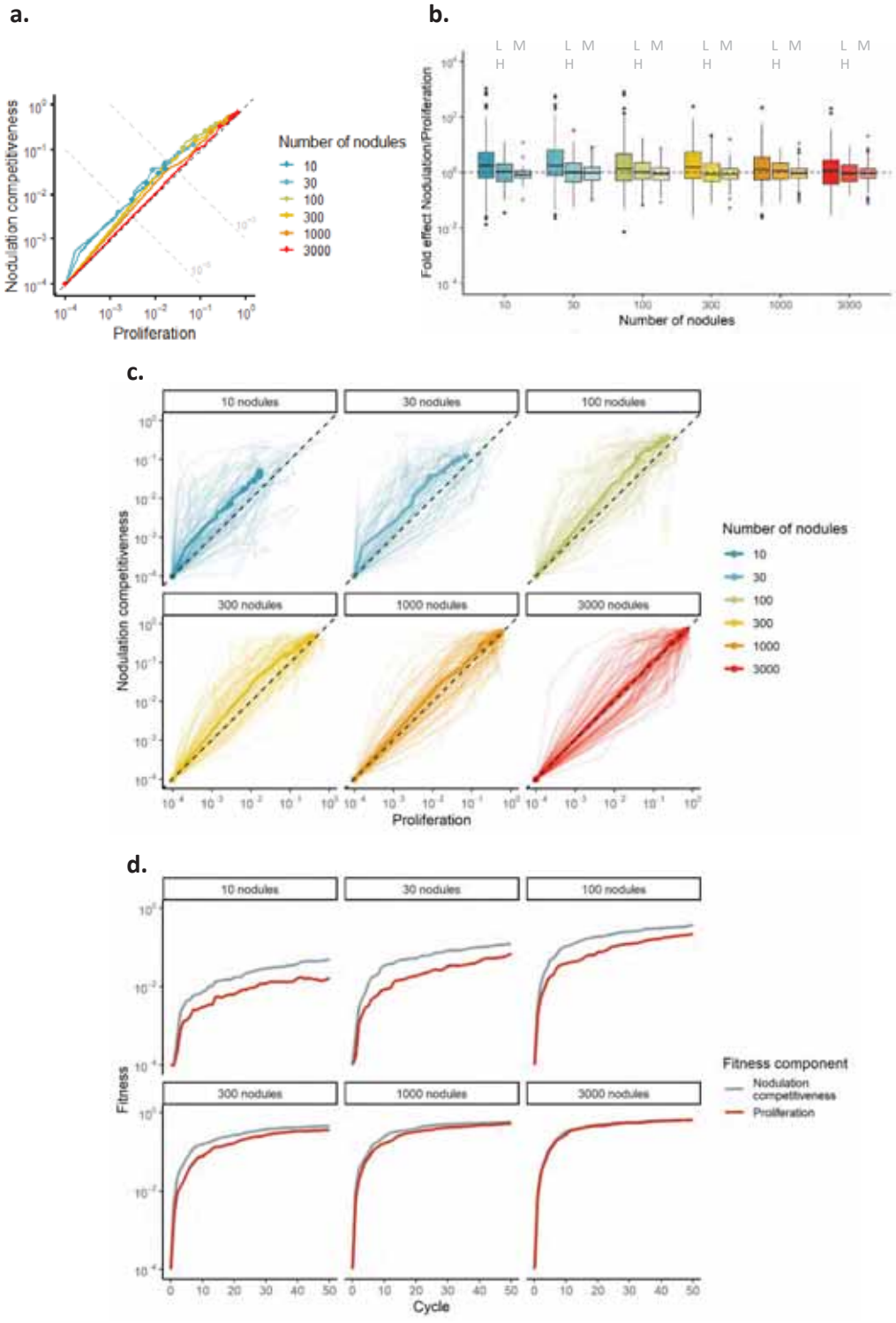
S2.12 Fig. Effect of the fitness of the ancestor on the relative strength of selection for nodulation competitiveness and proliferation, with a higher probability of beneficial mutations. **a,b**, as in Fig 6c,d (main text). **c,d**, as in S8 Fig, with fitness trajectories shown for various combinations of nodulation competitiveness/proliferations fitness values in the ancestor ($10^{-2}/10^{-4}$), ($10^{-2}/10^{-4}$), ($10^{-4}/10^{-4}$), ($10^{-4}/10^{-3}$), and ($10^{-4}/10^{-2}$). Evolutionary parameters used were: k = l = 3 and m = 0.



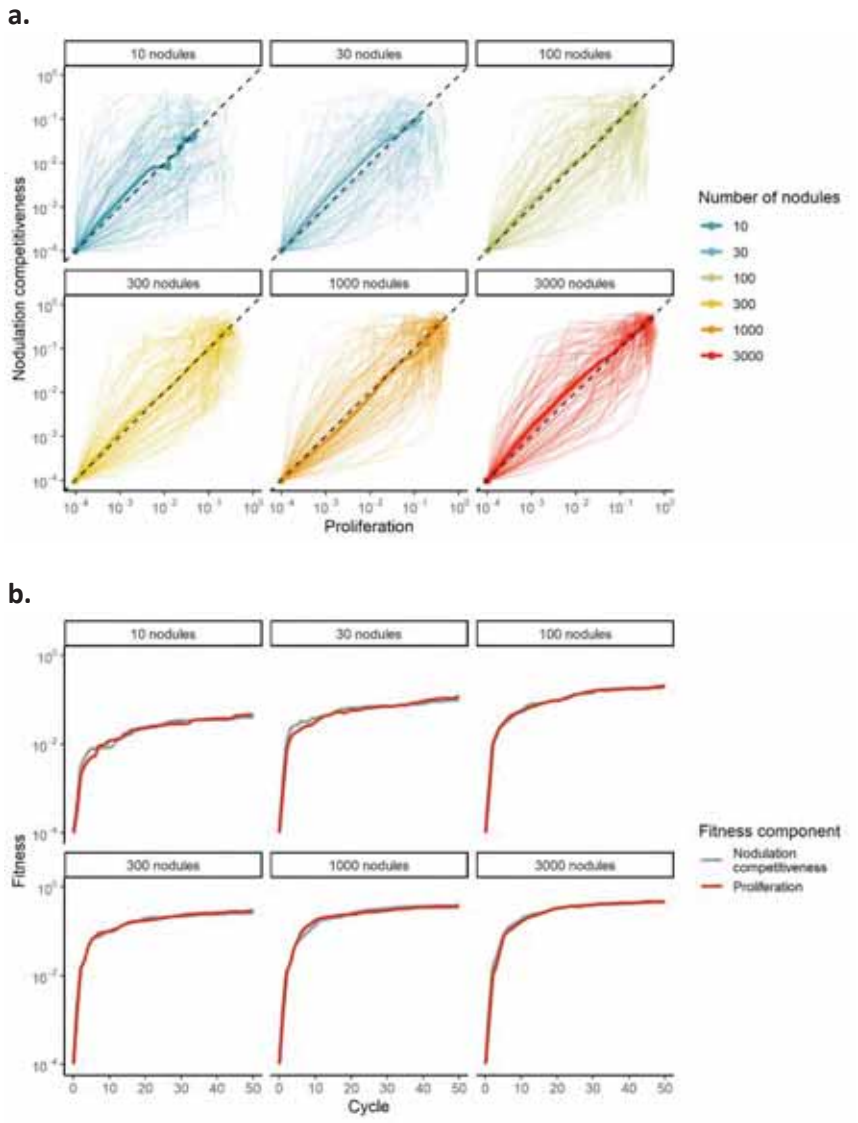
S2.13 Fig. Effect of the fitness of the ancestor on the relative strength of selection for nodulation competitiveness and proliferation, with a lower probability of beneficial mutations**. a,b**, as in Fig 6c,d (main text). **c,d**, as in S8 Fig, with fitness trajectories shown for various combinations of nodulation competitiveness/proliferations fitness values in the ancestor ($10^{-2}/10^{-4}$), ($10^{-2}/10^{-4}$), ($10^{-4}/10^{-4}$), ($10^{-4}/10^{-3}$), and ($10^{-4}/10^{-2}$). Evolutionary parameters used were: k = l = 20 and m = 0.



S2.14 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation under weak partial pleiotropy. **a,b**, as in Fig 6a,b (main text). **c,d**, as in S8 Fig. Evolutionary parameters used were: k = l = 6 and m = 4.



S2.15 Fig. Effect of the nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation under strong partial pleiotropy. **a,b**, as in Fig 6a,b (main text). **c,d**, as in S8 Fig. Evolutionary parameters used were: k = l = 2 and m = 8.



S2.16 Fig. Effect of the chronology of symbiotic events and the size of nodulation bottleneck on the relative strength of selection for nodulation competitiveness and proliferation. Same data as in Fig 6e,f (main text). **a,b**, as in S8 Fig. Evolutionary parameters used were: k = l = 10 and m = 0.

S2.1 Table. Number of nodules harvested along the evolutionary cycles

Cycle	Inoculum size for 30 plants in lineage B[a]	Number of harvested nodules in lineage B[b]	Inoculum size for 30 plants in lineage F[a]	Number of harvested nodules in lineage F[b]	Inoculum size for 30 plants in lineage G[a]	Number of harvested nodules in lineage G[b]	Inoculum size for 30 plants in lineage K[a]	Number of harvested nodules in lineage K[b]	Inoculum size for 30 plants in lineage M[a]	Number of harvested nodules in lineage M[b]
1	3.00E+07	82	3.00E+07	168	3.00E+07	196	3.00E+07	243	3.00E+07	35
2	2.00E+05	134	8.70E+05	125	1.30E+06	75	7.50E+05	66	5.20E+04	26
3	1.35E+06	87	3.70E+06	70	2.20E+06	112	3.70E+06	121	5.50E+08	19
4	1.00E+07	95	5.05E+08	178	1.20E+07	46	6.50E+06	59	5.00E+08	68
5	3.20E+06	107	8.00E+06	280	3.06E+08	49	2.00E+08	158	4.10E+05	103
6	1.00E+05	38	8.50E+07	308	6.00E+06	141	1.90E+07	298	3.50E+05	84
7	1.00E+07	37	5.00E+06	170	1.00E+07	51	2.60E+07	203	1.00E+07	34
8	1.00E+07	56	1.44E+07	238	4.00E+08	79	6.00E+06	275	1.00E+07	98
9	2.50E+07	169	1.52E+07	421	5.00E+09	161	4.43E+07	227	1.60E+06	59
10	4.50E+07	143	1.48E+07	200	4.00E+07	188	1.30E+07	287	1.10E+07	132
11	2.30E+07	165	6.70E+06	202	3.70E+07	158	1.00E+07	295	1.50E+06	148
12	2.30E+07	175	5.25E+06	228	5.00E+06	170	3.55E+06	216	2.60E+05	131
13	1.30E+07	161	1.00E+07	117	3.10E+07	309	1.00E+07	193	6.00E+05	80
14	2.75E+07	163	6.50E+06	230	1.80E+08	224	1.16E+07	130	6.65E+05	74
15	5.60E+07	169	4.40E+06	264	2.30E+07	272	8.30E+06	200	1.00E+08	58
16	3.35E+07	241	1.50E+07	189	1.55E+07	230	8.00E+06	116	1.00E+08	71
17	1.50E+07	117	1.50E+07	256	1.50E+07	218	1.50E+07	379	1.50E+07	143
18	2.68E+07	275	1.92E+07	110	5.98E+07	226	2.26E+07	165	1.49E+07	151
19	6.60E+07	201	3.74E+07	81	6.45E+07	153	4.29E+07	112	1.47E+07	31
20	7.35E+07	127	6.08E+07	220	1.22E+08	97	2.79E+07	187	5.85E+06	46
21	8.40E+07	167	5.85E+07	241	8.70E+07	163	2.55E+07	211	2.76E+06	66
22	6.75E+07	253	7.05E+07	98	6.60E+07	245	3.46E+07	164	1.79E+07	96
23	6.15E+07	134	5.79E+07	144	6.90E+07	147	5.52E+07	129	3.47E+07	98
24	3.08E+07	317	1.15E+08	122	8.70E+07	370	5.96E+07	197	3.98E+07	144
25	8.93E+07	182	1.17E+08	153	9.15E+07	154	7.50E+07	173	3.86E+07	85
26	5.48E+07	102	1.33E+08	97	1.10E+08	177	6.75E+07	293	2.36E+07	18
27	4.45E+07	106	7.58E+07	188	1.01E+08	216	1.14E+08	164	5.07E+06	76

28	7.35E+07	124	4.61E+07	114	1.24E+08	147	4.58E+07	169	3.27E+07	10
29	3.87E+07	97	1.06E+08	272	9.60E+07	156	6.04E+07	312	9.90E+06	45
30	4.65E+07	236	6.50E+07	186	1.29E+08	260	1.20E+06	419	8.10E+07	102
31	1.88E+07	226	1.73E+08	145	5.40E+07	359	1.61E+08	363	2.55E+07	101
32	1.31E+08	216	8.09E+07	320	1.40E+08	165	8.85E+07	242	3.82E+07	68
33	4.40E+07	261	6.60E+07	207	7.95E+07	176	7.65E+07	281	2.38E+07	109
34	4.10E+07	252	7.50E+07	290	9.15E+07	208	4.79E+07	242	2.15E+07	52
35	5.10E+07	179	6.15E+07	262	1.31E+08	189	6.00E+07	303	1.75E+07	56

[a] The inoculum size correspond to 1/10 of the total number of bacteria recovered from the crush of all nodules formed in the previous cycle,

S2.2 Table. Summary of detected mutations in the five lineages

Lineage	Cycle	Estimated bacterial generations[a]	Number of mutations (freq >5%)[b]	Number of fixed mutations	Number of new mutations	Total number of mutations/lineage (freq >5%)[b]	Total number of mutations/lineage (freq>30%)[c]	Total number of fixed mutations/lineage	Number of cohorts/lineage	Mean number of mutations/cohort
B	1	25	27	0	27	710	133	28	17	7.8
	3	75	50	1	38					
	5	125	77	1	43					
	7	175	82	2	41					
	9	225	57	11	21					
	11	275	64	28	15					
	13	325	73	28	19					
	15	375	74	28	19					
	17	425	86	28	32					
	19	475	105	28	31					
	21	525	133	28	61					
	23	575	160	28	61					
	25	625	152	28	55					
	27	675	183	28	47					
	29	725	223	28	78					
	31	775	209	28	38					
	33	825	230	28	48					
	35	875	195	28	36					
F	1	25	7	0	7	868	202	114	17	11.9
	3	75	58	0	54					
	5	125	55	0	8					
	7	175	62	0	17					
	9	225	73	0	22					
	11	275	107	0	46					
	13	325	150	25	107					
	15	375	134	31	41					
	17	425	203	31	106					
	19	475	235	31	74					
	21	525	218	74	49					

	23	575	217	74	34					
	25	625	191	95	44					
	27	675	202	103	37					
	29	725	228	111	32					
	31	775	275	114	61					
	33	825	310	114	81					
	35	875	306	114	48					
G	1	25	11	0	11	950	155	50	21	7.4
	3	75	69	1	64					
	5	125	86	1	43					
	7	175	83	1	32					
	9	225	118	1	37					
	11	275	139	1	41					
	13	325	69	1	14					
	15	375	71	30	14					
	17	425	96	30	25					
	19	475	125	30	50					
	21	525	165	30	77					
	23	575	205	30	59					
	25	625	227	33	81					
	27	675	174	39	74					
	29	725	226	39	68					
	31	775	272	39	59					
	33	825	336	50	117					
	35	875	309	50	84					
K	1	25	12	0	12	1204	233	147	24	9.7
	3	75	117	0	112					
	5	125	143	0	83					
	7	175	183	0	56					
	9	225	166	47	60					
	11	275	186	47	58					
	13	325	175	47	45					
	15	375	179	47	38					

	17	425	181	89	29					
	19	475	214	89	51					
	21	525	270	89	88					
	23	575	288	91	61					
	25	625	290	91	58					
	27	675	324	91	40					
	29	725	359	100	88					
	31	775	302	100	48					
	33	825	360	133	128					
	35	875	419	147	149					
M	1	25	17	0	17	382	96	33	22	4.4
	3	75	17	5	11					
	5	125	38	5	25					
	7	175	37	9	25					
	9	225	61	9	34					
	11	275	57	9	13					
	13	325	44	14	5					
	15	375	49	14	7					
	17	425	60	14	13					
	19	475	81	14	27					
	21	525	73	33	16					
	23	575	85	33	21					
	25	625	109	33	30					
	27	675	119	33	40					
	29	725	107	33	28					
	31	775	100	33	12					
	33	825	105	33	15					
	35	875	144	33	43					

[a]The number of bacterial generations was estimated according to (Remigi et al. 2014). [b] Number of mutations observed in a frequency > 5% in the evolved populations. [c]Number of mutations observed in a frequency

S2.3 Table. Global analysis of mutational cohorts

Lineage	Cohort	Cycle of appearance (app)[a]	Cycle of fixation (fix)[b]	Cycle of extinction (ext)[c]	Time to fixation (ttf)[d]	Initial speed of frequency increase (sup)[e]	Maximum frequency (maxF)[f]	Cycle of maximum frequency (Cfmax)[g]	Cohort fixation[h]	Cohort status[i]	Number of mutations in the cohort
B	1	1	3	0	2	0.124	1	3	+	Fixed	1
B	2	1	7	0	6	0.048	1	7	+	Fixed	1
B	3	1	9	0	8	0.033	1	9	+	Fixed	4
B	4	1	Not-fixed	9	0	0.042	0.340	5	-	Extinct	5
B	5	5	11	0	6	0.082	1	11	+	Fixed	16
B	6	7	11	0	4	0.049	1	11	+	Fixed	6
B	7	7	Not-fixed	0	0	0.052	0.929	35	-	Unknown	2
B	8	7	Not-fixed	0	0	0.024	0.773	27	-	Unknown	3
B	9	9	Not-fixed	25	0	0.033	0.323	15	-	Extinct	3
B	10	11	Not-fixed	0	0	0.017	0.922	35	-	Unknown	18
B	11	13	Not-fixed	0	0	0.033	0.511	27	-	Unknown	16
B	12	25	Not-fixed	0	0	0.033	0.389	33	-	Unknown	1
B	13	25	Not-fixed	0	0	0.017	0.663	35	-	Unknown	5
B	14	25	Not-fixed	0	0	0.028	0.384	33	-	Unknown	14
B	15	27	Not-fixed	0	0	0.021	0.317	33	-	Unknown	5
B	16	29	Not-fixed	0	0	0.016	0.614	35	-	Unknown	28
B	17	33	Not-fixed	0	0	0.048	0.373	35	-	Unknown	5
F	1	3	Not-fixed	13	0	0.080	0.740	3	-	Extinct	4
F	2	3	Not-fixed	13	0	0.048	0.369	3	-	Extinct	12
F	3	3	15	0	12	0.022	1	15	+	Fixed	14
F	4	5	13	0	8	0.052	1	13	+	Fixed	12
F	5	9	15	0	6	0.022	1	15	+	Fixed	5
F	6	13	Not-fixed	19	0	0.039	0.378	15	-	Extinct	12
F	7	9	21	0	12	0.050	1	21	+	Fixed	34
F	8	13	21	0	8	0.039	1	21	+	Fixed	9
F	9	15	Not-fixed	21	0	0.057	0.477	17	-	Extinct	26
F	10	17	29	0	12	0.058	1	29	+	Fixed	23

F	11	19	Not-fixed	31	0	0.034	0.408	23	-	Extinct	5
F	12	17	Not-fixed	27	0	0.036	0.272	21	-	Extinct	7
F	13	19	31	0	12	0.031	1	31	+	Fixed	15
F	14	21	Not-fixed	0	0	0.023	0.454	27	-	Unknown	1
F	15	19	31	0	12	0.045	1	31	+	Fixed	2
F	16	23	Not-fixed	0	0	0.032	0.407	35	-	Unknown	14
F	17	27	Not-fixed	0	0	0.022	0.317	35	-	Unknown	7
G	1	1	3	0	2	0.095	1	3	+	Fixed	1
G	2	3	15	0	12	0.049	1	15	+	Fixed	3
G	3	3	Not-fixed	15	0	0.024	0.408	5	-	Extinct	34
G	4	5	15	0	10	0.061	1	15	+	Fixed	5
G	5	5	Not-fixed	9	0	0.037	0.389	7	-	Extinct	1
G	6	7	15	0	8	0.085	1	15	+	Fixed	21
G	7	9	27	0	18	0.026	1	27	+	Fixed	5
G	8	11	27	0	16	0.021	1	27	+	Fixed	4
G	9	11	Not-fixed	27	0	0.029	0.694	17	-	Extinct	1
G	10	13	Not-fixed	27	0	0.040	0.352	17	-	Extinct	4
G	11	13	33	0	20	0.017	1	33	+	Fixed	1
G	12	17	Not-fixed	27	0	0.030	0.357	19	-	Extinct	2
G	13	19	33	0	14	0.050	1	33	+	Fixed	8
G	14	19	Not-fixed	0	0	0.039	0.389	25	-	Unknown	7
G	15	21	Not-fixed	0	0	0.030	0.978	35	-	Unknown	16
G	16	23	Not-fixed	0	0	0.024	0.720	35	-	Unknown	9
G	17	25	Not-fixed	0	0	0.034	0.528	33	-	Unknown	4
G	18	27	Not-fixed	0	0	0.017	0.517	35	-	Unknown	7
G	19	27	Not-fixed	0	0	0.037	0.365	33	-	Unknown	5
G	20	29	Not-fixed	31	0	0.046	0.344	29	-	Extinct	1
G	21	29	Not-fixed	0	0	0.029	0.424	35	-	Unknown	16
K	1	3	9	0	6	0.054	1	9	+	Fixed	46
K	2	3	Not-fixed	9	0	0.043	0.313	3	-	Extinct	2
K	3	3	9	0	6	0.031	1	9	+	Fixed	1
K	4	5	17	0	12	0.056	1	17	+	Fixed	19
K	5	5	17	0	12	0.048	1	17	+	Fixed	6

K	6	7	Not-fixed	17	0	0.033	0.569	9	-	Extinct	1
K	7	7	Not-fixed	25	0	0.026	0.393	11	-	Extinct	11
K	8	7	Not-fixed	17	0	0.056	0.462	9	-	Extinct	19
K	9	9	17	0	8	0.054	1	17	+	Fixed	17
K	10	11	23	0	12	0.041	1	23	+	Fixed	2
K	11	11	Not-fixed	29	0	0.019	0.421	19	-	Extinct	4
K	12	13	29	0	16	0.040	1	29	+	Fixed	9
K	13	13	Not-fixed	33	0	0.025	0.463	21	-	Extinct	1
K	14	13	33	0	20	0.036	1	33	+	Fixed	10
K	15	13	Not-fixed	31	0	0.028	0.356	21	-	Extinct	6
K	16	15	Not-fixed	29	0	0.030	0.301	17	-	Extinct	1
K	17	19	33	0	14	0.019	1	33	+	Fixed	23
K	18	21	Not-fixed	31	0	0.039	0.367	23	-	Extinct	1
K	19	21	Not-fixed	35	0	0.034	0.487	25	-	Extinct	14
K	20	21	35	0	14	0.019	1	35	+	Fixed	14
K	21	25	Not-fixed	0	0	0.017	0.805	33	-	Unknown	3
K	22	27	Not-fixed	31	0	0.066	0.592	27	-	Extinct	1
K	23	27	Not-fixed	0	0	0.038	0.505	35	-	Unknown	20
K	24	29	Not-fixed	0	0	0.037	0.379	33	-	Unknown	2
M	1	3	3	0	0	0.151	1	3	+	Fixed	5
M	2	3	Not-fixed	5	0	0.060	0.509	3	-	Extinct	1
M	3	5	7	0	2	0.078	1	7	+	Fixed	1
M	4	7	7	0	0	0.151	1	7	+	Fixed	3
M	5	7	13	0	6	0.082	1	13	+	Fixed	4
M	6	7	13	0	6	0.057	1	13	+	Fixed	1
M	7	9	Not-fixed	21	0	0.057	0.981	15	-	Extinct	4
M	8	9	Not-fixed	11	0	0.043	0.310	9	-	Extinct	1
M	9	9	21	0	12	0.016	1	21	+	Fixed	12
M	10	11	Not-fixed	21	0	0.056	0.984	15	-	Extinct	4
M	11	13	Not-fixed	15	0	0.051	0.399	13	-	Extinct	1
M	12	13	21	0	8	0.080	1	21	+	Fixed	7
M	13	13	Not-fixed	21	0	0.053	0.427	15	-	Extinct	3
M	14	17	Not-fixed	0	0	0.029	0.966	31	-	Unknown	2

M	15	17	Not-fixed	0	0	0.020	0.432	27	-	Unknown	7
M	16	19	Not-fixed	29	0	0.042	0.303	19	-	Extinct	3
M	17	23	Not-fixed	0	0	0.016	0.960	33	-	Unknown	7
M	18	23	Not-fixed	29	0	0.044	0.323	27	-	Extinct	2
M	19	25	Not-fixed	0	0	0.018	0.739	31	-	Unknown	17
M	20	27	Not-fixed	0	0	0.030	0.383	31	-	Unknown	8
M	21	29	Not-fixed	33	0	0.043	0.306	29	-	Extinct	1
M	22	31	Not-fixed	0	0	0.018	0.346	35	-	Unknown	2

[a]app, cycle of appearance of the cohort. [b]fix, cycle when the cohort reaches a frequency of 100% . [c] ext, cycle when the cohort is extinct (reaches a frequency of 0%). [d]ttf, number of cycles between the appearance and the fixation of the cohort. [e]sup, initial speed of frequency increase (adapted from Lang et al., 2013). [f]maxF, maximum frequency reached by the cohort. [g]Cfmax, cycle when the maximum frequency of the cohort is reached. [h]Cohort fixation, fixation status of the cohort after 35 cycles of experimental evolution. [i]Cohort status after 35 cycles of experimental evolution.

S2.4 Table. Neutral and deleterious reconstructed mutations

Lineage	Gene ID[a]	Gene name	Mutation	Type of mutation[b]	Description	Cohort	Mean CI for *in planta* fitness[c]	Symbiotic effect	G score[d]
B	pRALTA_0449		D184N	Nonsyn	Cytochrome P450-terp	CoB3	0.7	neutral	8.02
	up[66]-RSp0211		G/A	Intergenic	Hypothetical protein	CoB6	**0.4**	deleterious	nd
	RSp1071		Q295*	Nonsense	Putative hemagglutinin-related protein	CoB6	**0.3**	deleterious	-0.85
	pRALTA_0288		V193D	Nonsyn	Putative multi-sensor signal transduction histidine kinase	CoB6	0.9	neutral	20.41
	pRALTA_0291		A69A	Syn	Putative aldose 1-epimerase/galactose mutarotase	CoB6	**0.7**	deleterious	7.89
	RSc0711	*nusB*	L69L	Syn	Transcription antitermination protein NusB	CoB7	0.4	neutral	1.39
	pRALTA_0615		L93L	Syn	Integrase fragment	CoB10	1.5	neutral	3.43
G	RSc1962		S9F	Nonsyn	Putative transmembrane protein	CoG2	1.2	neutral	5.07
	RSp1439	*cysE1*	P248L	Nonsyn	Serine acetyltransferase protein	CoG7	1.5	neutral	1.43
	RSc2277		F16F	Syn	Putative transporter protein	CoG8	0.9	neutral	36.95
	RSc2326		G116S	Nonsyn	Putative lipoprotein	CoG11	0.8	neutral	1.59
	RSc0672		F948F	Syn	Putative sensor histidine kinase and response regulator protein	CoG13	1.1	neutral	0.28
	RSc0537	*fadE*	A378A	Syn	Acyl-CoA dehydrogenase	CoG15	1.2	neutral	4.15
	RSc1311	*dnaB*	S21S	Syn	Replicative DNA helicase	CoG15	**0.5**	deleterious	3.41
	RSc2383		V926A	Nonsyn	Putative signal peptide protein	CoG15	0.8	neutral	1.50
	RSp0208		I88I	Syn	Hypothetical protein	CoG15	0.9	neutral	15.12

[a] up[x], intergenic mutations located x nucleotides upstream the gene indicated. [b]Syn, synonymous mutations. Nonsyn, non synonymous mutations. [c]Mean values of competitive indexes obtained from 3 to 4 independent experiments. Figures in bold are statistically different from 1 ($P<0.05$, t-test with Benjamini-Hochberg correction). [d]G scores were calculated as proposed by Tenaillon et al. [23]. nd, G scores were not determined for mutations in pseudogenes or intergenic regions.

S2.5 Table. Frequencies of usage of codons modified by adaptive synonymous mutations

Synonymous adaptive mutation	Wild-type codon	Mutated codon	Frequency of usage of wild-type codon	Frequency of usage of mutated codon
RSp0595-P132P	CCG	CCA	34.7	2.2
rhlE1-L368L	CTG	TTG	73.3	6.3
RSc0243-L213L	CTC	CTT	19.7	3
RSp1116-I32I	ATC	ATT	38.4	4.4
pRALTA_0552-I100I	ATC	ATA	38.4	0.6
RSc1982-I100I	ATC	ATT	38.4	4.4
RSc0672-F948F	TTC	TTT	29.2	4.3
RSp0574-L15L	CTC	CTT	19.7	3
RSp0587-P84P	CCT	CCA	2.1	2.2
RSp0759-F91F	TTC	TTT	29.2	4.3

S2.6 Table. Strains, plasmids, oligonucleotides and softwares used in this study.

Resource type	Designation	Identifier	Additional information/Common name/5'-3' sequence	Reference/source
Cupriavidus taiwanensis	Wild-type strain isolated from *Mimosa pudica* in Taiwan	LMG19424		Chen et al., 2001
Cupriavidus taiwanensis	LMG19424 derivative resistant to StrR	CBM832		M. Hynes
Ralstonia solanacearum	Non-nodulating chimeric ancestor, TriR	CBM124	GMI1000 pRalta::TriR	Marchetti et al., 2010
Ralstonia solanacearum	Nodulating ancestor mutated in *hrpG* , TriR, GenR	CBM212	CBM212	Marchetti et al., 2010
Ralstonia solanacearum	Nodulating ancestor mutated in *hrpG* , TriR, GenR	CBM349	CBM349	Marchetti et al., 2010
Ralstonia solanacearum	Nodulating ancestor mutated in *hrcV* , TriR, GenR	CBM356	CBM356	Marchetti et al., 2010
Ralstonia solanacearum	Evolved clone from line B, cycle 16, TriR, GenR	CBM1492	Clone B16	Marchetti et al., 2014
Ralstonia solanacearum	Evolved clone from line F, cycle 16, TriR, GenR	CBM2182	Clone F16	Marchetti et al., 2017
Ralstonia solanacearum	Evolved clone from line G, cycle 16, TriR, GenR	CBM1494	Clone G16	Marchetti et al., 2014
Ralstonia solanacearum	Evolved clone from line K, cycle 16, TriR, GenR	CBM2184	Clone K16	Marchetti et al., 2017
Ralstonia solanacearum	Evolved clone from line M, cycle 16, TriR, GenR	CBM1497	Clone M16	Guan et al., 2013
Ralstonia solanacearum	Evolved clone from line S, cycle 16, TriR, GenR	CBM1752	Clone S16	Guan et al., 2013
Ralstonia solanacearum	Evolved clone from line B, cycle 35, TriR, GenR	RCM2458	Clone B35	This study
Ralstonia solanacearum	Evolved clone from line F, cycle 35, TriR, GenR	RCM2475	Clone F35	This study
Ralstonia solanacearum	Evolved clone from line G, cycle 35, TriR, GenR	RCM2460	Clone G35	This study
Ralstonia solanacearum	Evolved clone from line K, cycle 35, TriR, GenR	RCM2476	Clone K35	This study
Ralstonia solanacearum	Evolved clone from line M, cycle 35, TriR, GenR	RCM2462	Clone M35	This study
Ralstonia solanacearum	Evolved clone from line S, cycle 35, TriR, GenR	RCM2477	Clone S35	This study
Ralstonia solanacearum	Evolved clone from line B, cycle 1, TriR, GenR	CBM418	Clone B1	Marchetti et al., 2010
Ralstonia solanacearum	Evolved clone from line B, cycle 3, TriR, GenR	CBM510	Clone B3	Marchetti et al., 2010
Ralstonia solanacearum	Evolved clone from line B, cycle 20, TriR, GenR	RCM2046	Clone B20	This study
Ralstonia solanacearum	Evolved clone from line G, cycle 1, TriR, GenR	CBM420	Clone G1	Marchetti et al., 2010
Ralstonia solanacearum	Evolved clone from line G, cycle 2, TriR, GenR	CBM465	Clone G2	Marchetti et al., 2010
Ralstonia solanacearum	Evolved clone from line G, cycle 4, TriR, GenR	CBM557	Clone G4	Marchetti et al., 2010
Ralstonia solanacearum	Evolved clone from line G, cycle 7, TriR, GenR	CBM688	Clone G7	Marchetti et al., 2010
Ralstonia solanacearum	Evolved clone from line G, cycle 11, TriR, GenR	CBM1281	Clone G11	Marchetti et al., 2010
Ralstonia solanacearum	Evolved clone from line G, cycle 25, TriR, GenR	RCM2162	Clone G25	This study
Ralstonia solanacearum	Nodulating ancestor CBM212 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM1833	CBM212 IGglmS:P*psbA* -GFP	Marchetti et al., 2017
Ralstonia solanacearum	Nodulating ancestor CBM212 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM1844	CBM212 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B1 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM1961	Clone B1 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B1 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM1962	Clone B1 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B3 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM3343	Clone B3 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B3 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM3040	Clone B3 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B20 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM3037	Clone B20 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B20 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM3349	Clone B20 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G1 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM1818	Clone G1 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G1 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM1819	Clone G1 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G2 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM1824	Clone G2 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G2 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM1827	Clone G2 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G4 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM1958	Clone G4 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G4 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM3350	Clone G4 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G7 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM2996	Clone G7 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G7 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM3613	Clone G7 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G11 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM1979	Clone G11 IGglmS:P*psbA* -GFP	This study

Ralstonia solanacearum	Evolved clone G11 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM2997	Clone G11 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G25 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	RCM3000	Clone G25 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G25 carrying a P*psbA* -mCherry fusion downstream glmS, TriR, KanR	RCM2999	Clone G25 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	CBM212 Rsp0595-P132P carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS, TriR, KanR	RCM3130	CBM212 Rsp0595-P132P IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	CBM212 RSc1598-P455L carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR, KanR	RCM2166	CBM212 RSc1598-P455L IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B1 pRALTA_0449-D184N carrying a PpsbA-GFP fusion and a KanR cassette downstream	RCM3079	B1 pRALTA_0449-D184N IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B1 Rsp0788-E225K carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3131	B1 Rsp0788-E225K IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B1 RSp1590-S153L carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3280	B1 RSp1590-S153L IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B1 RSp1417 -S417L carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3282	B1 RSp1417-S417L IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B3 up^{66}-RSp0211 carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS. TriR. KanR	RCM3153	B3 up^{66}-RSp0211 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B3 pRALTA_0291 - A69A carrying a P*psbA* -GFP fusion and a KanR cassette downstream	RCM3154	B3 pRALTA_0291-A69A IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B3 up^{115}-RSc0965 carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS. TriR. KanR	RCM3157	B3 up^{115}-RSc0965 IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B3 pRalta_0288-V193D carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3069	Clone B3 pRalta_0288-V193D IGglmS:P*psbA* -mcherry	This study
Ralstonia solanacearum	Evolved clone B3 RSp1071-Q295* carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR, KanR	RCM3210	B3 RSp1071-Q295* IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B3 *nusB* -L69L carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR, KanR	RCM3211	B3 *nusB* -L69L IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone B3 *rhlE1* -L368L carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3276	B3 *rhlE1* -L368L-IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B20 pRALTA_0552 carrying a PpsbA-mCherry fusion and a KanR cassette downstream glmS,	RCM3277	B20 pRALTA_0552-I100I IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B20 pRALTA_0615-L93L carrying a PpsbA-mCherry fusion and a KanR cassette downstream glmS,	RCM3312	B20 pRALTA_0615-L93L IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B20 up^{8}-RSp1469 carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS.	RCM3376	B20 up^{8}-RSp1469 IGglmS:P*psbA* -mcherry	This study
Ralstonia solanacearum	Evolved clone B20 pRALTA_0230-G242E carrying a PpsbA-mCherry fusion and a KanR cassette downstream	RCM3520	B20 pRALTA_0230-G242E IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone B20 RSc0243-L213L carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3521	B20 RSc0243-L213L IGglmS:P*psbA* -mcherry	This study
Ralstonia solanacearum	Evolved clone B20 RSp1116-I32I carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3522	B20 RSp1116-I32I IGglmS:P*psbA* -mcherry	This study
Ralstonia solanacearum	Evolved clone G1 *efpR* -E66K carrying a KanR cassette downstream glmS, TriR, KanR	RCM1867	G1 *efpR* -E66K IGglmS:KanR	Capela et al., 2017
Ralstonia solanacearum	Evolved clone G2 RSc1982 - I100I carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3156	G2 RSc1982-I100I IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G2 RSc1962 - S9F carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR, KanR	RCM3155	G2 RSc1962-S9F IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G4 *ppiB* -Q112* carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3084	G4 *ppiB* -Q112* IGglmS:P*psbA* -mcherry	This study
Ralstonia solanacearum	Evolved clone G4 RSc2277-V321G carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3083	G4 RSc2277-V321G IGglmS:P*psbA* -mcherry	This study
Ralstonia solanacearum	Evolved clone G7 up^{180}-RSp0637 carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS.	RCM3214	G7 up^{180}-RSp0637 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G7 *cysE1* -P248L carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3500	G7 *cysE1* -P248L IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G7 up^{85}-*gloA* carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS. TriR. KanR	RCM3310	G7 up^{85}-*gloA* IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G11 RSc2277-F16F carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR, KanR	RCM3556	G11 RSc2277-F16F IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G11 RSc2326 -G116S carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR,	RCM3278	G11 RSc2326-G116S IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G11 RSc0194-S102F carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3111	G11 RSc0194-S102F IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G11 Rsp0741 (tssJ)-I157N carrying a P*psbA* -mCherry fusion and a KanR cassette downstream	RCM3132	G11 Rsp0741-I157N IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G11 *phbB* -S38F carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3212	G11 *phbB* -S38F IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G11RSc0672-F948F carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR, KanR	RCM3313	G11 RSc0672-F948F IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G25 up^{116}-RSp0780 carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS.	RCM3088	G25 up^{116}-RSp0780 IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G25 RSp0759-P92S carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3087	G25 RSp0759-P92S IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G25 pRALTA_0398-W719* carrying a PpsbA-mCherry fusion and a KanR cassette downstream	RCM3216	G25 pRALTA_0398-W719* IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G25 RSp0769 - F416S carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3158	G25 RSp0769-F416S IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G25 RSp0587- P84P carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3309	G25 RSp0587-P84P IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G25 RSp0574-L15L carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3219	G25 RSp0574-L15L IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G25 *dnaB* -S21S carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR, KanR	RCM3208	G25 *dnaB* -S21S IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G25 RSp0208 - G92Q carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR,	RCM3281	G25 RSp0208-G92Q IGglmS:P*psbA* -GFP	This study
Ralstonia solanacearum	Evolved clone G25RSp1422 - G1298E carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3279	G25 RSp1422-G1298E IGglmS:P*psbA* -mCherry	This study
Ralstonia solanacearum	Evolved clone G25 *fadE* -A378A carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS,	RCM3207	G25 *fadE* -A378A IGglmS:P*psbA* -mCherry	This study

Ralstonia solanacearum	Evolved clone G25 RSc2383-V926A carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR,	RCM3375	G25 RSc2383-V926A IGglmS:P*psbA* -GFP	This study
Plasmid	Symbiotic plasmid of LMG19424 (0.5 Mb), likely auto-transferable	pRalta		Amadou, et al. 2008
Plasmid	Plasmid for *R. solanacearum* chromosomal integration of the constitutive *psbA* promoter fused to GFPuv into	pRCK-P*ps* -GFP		Capela, et al. 2017
Plasmid	Plasmid for *R. solanacearum* chromosomal integration of the constitutive *psbA* promoter fused to mCherry into	pRCK-P*ps* -mCherry		Capela, et al. 2017
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 728875 in RSp0595 (CoB1)	oCBM4346	GGAGCAAGCTGACATGGGAA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 728875 in RSp0595 (CoB1)	oCBM4347	GGCTCGGACTTGATCTTGCG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1715630 in RSc1598 (CoB2)	oCBM2805	CGAGCTGATCGTCGGCAACC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1715630 in RSc1598 (CoB2)	oCBM2806	TTCAGGCGAGGTGACGGTCA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1781273 in RSp1417 (CoB3)	oCBM4356	GGCCACTTTGTAGGACTGCA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1781273 in RSp1417 (CoB3)	oCBM4357	TCGTGAATGACCGTCCGTTT	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1994246 in RSp1590 (CoB3)	oCBM4730	ACGATAGCGTCCTCCACCT	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1994246 in RSp1590 (CoB3)	oCBM4733	CGATTCGCCCGATCAAAACC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 992273 in RSp0788 (CoB3)	oCBM4351	GCTGGAGAAGGTCAACGTGA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 992273 in RSp0788 (CoB3)	oCBM4352	GACATCTCGATCGCCAGGTT	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 387383 in pRALTA_0449 (CoB3)	oCBM4366	AATCCCTTTCCGCCTGATCG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 387383 in pRALTA_0449 (CoB3)	oCBM4367	GTGTCTCACCACGGTCGAAT	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1013285 in up115-RSc0965 (CoB6)	oCBM2529	ACTGGGCCTATCTCAAGACG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1013285 in up115-RSc0965 (CoB6)	oCBM2530	CAACAGGTGGTCGATGAAGC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1353434 in RSp1071 (CoB6)	oCBM4381	CCAGAGCGCAAAGTAGTCGA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1353434 in RSp1071 (CoB6)	oCBM4382	TGACCCCGGAAGAGAAGGAA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 243268 in pRALTA_0288 (CoB6)	oCBM4371	CACTCCATCCGCACTCTTCG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 243268 in pRALTA_0288 (CoB6)	oCBM4372	CCGCTAGAGTTCACATCCGG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 247935 in pRALTA_0291 (CoB6)	oCBM4376	CTTGTGGGTGCTGATCGTCT	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 247935 in pRALTA_0291 (CoB6)	oCBM4377	TCGACGGCTTCCTTCAGATG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 259792 in up66-RSp0211 (CoB6)	oCBM2522	TCGGTCTTGCAGCTTGTAGA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 259792 in up66-RSp0211 (CoB6)	oCBM2523	AATCCACCTGTCAGTCCTGG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 583937 in rhlE1 (CoB7)	oCBM4734	CTTCATCACCACCTGCGAGT	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 583937 in rhlE1 (CoB7)	oCBM4735	CGCTATGGACACCCTCGTAC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 759825 in nusB (CoB7)	oCBM4396	TGATGACAGAACGAACGGCA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 759825 in nusB (CoB7)	oCBM4397	GATCCTGGTCTATCTGGCGC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1414257 in RSp1116 (CoB10)	oCBM4736	CTCAAACAACAACAGGGCGG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1414257 in RSp1116 (CoB10)	oCBM4737	ATGCTGGGGTGTTGATGTGG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the intergenic mutation position 1848419 in RSp1469 / RSp1470	oCBM3928	ATCGCGCAATGTCTGGATCA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the intergenic mutation position 1848419 in RSp1469 / RSp1470	oCBM3929	CCTTCACGTCCGGATAGTCG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 267735 in RSc0243 (CoB10)	oCBM3907	GAAGAGGTTGTAGGCGATGC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 267735 in RSc0243 (CoB10)	oCBM3908	CCAGTTGAGGTCGCACAAG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 197522 in pRALTA_0230 (CoB10)	oCBM4419	TCCCAACGTCGCTTTCATGA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 197522 in pRALTA_0230 (CoB10)	oCBM4420	CTCCATTTCCGCTGCGAATG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 463545 in pRALTA_0552 (CoB10)	oCBM4442	CAATCTGCGTCGTTGGTCAC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 463545 in pRALTA_0552 (CoB10)	oCBM4443	GTTTGAGGCCGAGTCCATGA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 502781 in pRALTA_0615 (CoB10)	oCBM4447	TACCCAACAGCGACAGTTCC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 502781 in pRALTA_0615 (CoB10)	oCBM4448	CCCTTCATCCTGCCGATCTC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1155501 in RSc1097 (CoG1)	oCBM4457	CCTCTACGATCACATCCCGG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1155501 in RSc1097 (CoG1)	oCBM4458	TTCGCCTTGATGAAGCCCTT	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2157029 in RSc1982 (CoG2)	oCBM4073	CGCCCTGGAAGTTGAAATCGC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2157029 in RSc1982 (CoG2)	oCBM4074	TTTCGGGATGGCAGTCGCAACC	This study

Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2141969 in RSc1962 (CoG2)	oCBM4068	CCGCTGGTTGATCCATTG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2141969 in RSc1962 (CoG2)	oCBM4069	ATGCCGCAGAGAAGGAAG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1222748 in ppiB (CoG4)	oCBM4086	TCCGGATACGCTGTCTTGATG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1222748 in ppiB (CoG4)	oCBM4087	TGAGCGCGTCTTGTTGACTTC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2474088 in RSc2277 (CoG4)	oCBM4077	CTCCTGCATCAGCACTTC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2474088 in RSc2277 (CoG4)	oCBM4078	GGATGGACGCGTACATTC	
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1808334 in cysE1 (CoG7)	oCBM4477	GCAGGAACGACGACCAGATC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1808334 in cysE1 (CoG7)	oCBM4478	CAGCAGATCGTTGATTCGCC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 558407 in gloA/RSc0521 (CoG7)	oCBM4482	CCATACCGAATACCTGGCCG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 558407 in gloA/RSc0521 (CoG7)	oCBM4483	CAACAAGCGCATCACCAACA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 777392 in RSp0637/RSp0638 (CoG7)	oCBM4492	GGATGGACGAGGTATAGCGC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 777392 in RSp0637/RSp0638 (CoG7)	oCBM4493	TCCGTTCCTTCTTCACCAGC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2473174 in RSc2277 (CoG8)	oCBM4742	GAAGCCATAGTACGGACCGG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2473174 in RSc2277 (CoG8)	oCBM4743	TCCAAACAAACGACACCCCA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2526677 in RSc2326 (CoG11)	oCBM4517	GCATGGTTGTTTCACTCGGG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 2526677 in RSc2326 (CoG11)	oCBM4518	CCATCATTCCTGCCGTCCAT	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1754626 in phbB (CoG13)	oCBM4543	AGTCCGCTGTATGGCTATGC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 1754626 in phbB (CoG13)	oCBM4544	CGGAAATCTCGTACCCCTCG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 215421 in RSc0194 (CoG13)	oCBM4541	CCTATGGCGGAAACCCTGTT	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 215421 in RSc0194 (CoG13)	oCBM4542	GGTGAGTTCGGGAAAGCTGA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 719867 in RSc0672 (CoG13)	oCBM4536	CATCATCCATCTTCCGCCCA	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 719867 in RSc0672 (CoG13)	oCBM4537	TTGCCCAGGTAGACGTTGAC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 931539 in RSp0741/tssJ (CoG13)	oCBM4524	CCAGGGCCTTGTAGTTCCAC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 931539 in RSp0741/tssJ (CoG13)	oCBM4525	GTATTTGCTGTGGTCGGTGC	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 345083 in pRALTA_0398 (CoG15)	oCBM3709	CCTGGTTTGGGGAAAGTGTG	This study
Sequence-based ragent- PCR primer	Amplification of 6 kb fragment around the mutation position 345083 in pRALTA_0398 (CoG15)	oCBM3710	CAACCCGCAGAATCAGAGTG	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 1395827 in dnaB (CoG15)	oCBM4289	TTCAAGATTTCGTCGCGCAT	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 1395827 in dnaB (CoG15)	oCBM4290	TTGGTGATCTTCTGCCCCAT	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 1789297 in RSp1422 (CoG15)	oCBM2500	CGCTCGATTACTACAAGGCG	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 1789297 in RSp1422 (CoG15)	oCBM2501	TTCCGTGGTTTCATCTCCCA	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 257773 in RSp0208 (CoG15)	oCBM4299	GCCATTCGTTGCCCTTCTTG	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 257773 in RSp0208 (CoG15)	oCBM4300	TCGAACAGGTGGTGATAGCG	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 2589097 in RSc2383 (CoG15)	oCBM3729	ATCCTGTACTTCCACACGCT	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 2589097 in RSc2383 (CoG15)	oCBM3730	CATCCAGCAGCACCAGTTC	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 579465 in fadE (CoG15)	oCBM4291	AAAGGCCCGCTCTGTCATC	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 579465 in fadE (CoG15)	oCBM4292	AGTTTGCACACACTCCCACA	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 713879 in RSp0574 (CoG15)	oCBM4293	AGTCACATGCTTCGCCTGAT	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 713879 in RSp0574 (CoG15)	oCBM4294	AAACTGAGCCTACCCAAGCC	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 722803 in RSp0587 (CoG15)	oCBM4297	TTGAGGTGCTGTTTGGGTGT	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 722803 in RSp0587 (CoG15)	oCBM4298	CCATGTCAGCTTGCTCCGTA	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 957425 in RSp0759 (CoG15)	oCBM3749	TGGATGGTTTTGCTTTGGGG	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 957425 in RSp0759 (CoG15)	oCBM3750	ATCAACATCGACGATCCGGA	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 973212 in RSp0769 (CoG15)	oCBM3759	CATCGGCCAGGAAGTGTTG	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 973212 in RSp0769 (CoG15)	oCBM3760	TTTTCGACGCGGACTGTTC	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 980536 in RSp0779/RSp0780 (CoG15)	oCBM3764	AGGACTATCGCTCTTTCGCA	This study
Sequence-based ragent- PCR primer	Amplification of 6kb fragment around mutation position 980536 in RSp0779/RSp0780 (CoG15)	oCBM3765	CGTAGCGCTCGTAATTCACC	This study

Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 728875 in RSp0595 (CoB1)	oCBM4682	AAGACAACGGCGAGCCCTA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 728875 in RSp0595 (CoB1)	oCBM4683	CCCTCCATCAACTTCGACCA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1781273 in RSp1417 (CoB3)	oCBM4763	AATGTATCCGCCTGGTCATC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1781273 in RSp1417 (CoB3)	oCBM4764	GGACGGTCACCTTGTTCTTC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1781273 in RSp1590 (CoB3)	oCBM4767	CAGATGAACTGCTCGCCTTC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1781273 in RSp1590 (CoB3)	oCBM4768	GCTACCTGTACCTGGCAACG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 387383 in pRALTA 0449 (CoB3)	oCBM4995	TGAAGGCCTGGTTCTGATCG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 387383 in pRALTA_0449 (CoB3)	oCBM4365	CCCAAAGACTTACTCAGACCG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 992273 in RSc0788 (CoB3)	oCBM4684	TACGAGCACCACCTCACG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 992273 in RSc0788 (CoB3)	oCBM4685	CCAGCATCTCCAGCATCAGT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1353434 in RSp1071 (CoB6)	oCBM4704	GTTGACCACCATGTTCTGCG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1353434 in RSp1071 (CoB6)	oCBM4705	ATCACGGATGGCGCGAAG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 243268 in pRALTA_0288 (CoB6)	oCBM5000	CAGCGTGATGAAGCCCGA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 243268 in pRALTA 0288 (CoB6)	oCBM4370	CGGGATGCTAGCAATTCGAG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 247935 in pRALTA 0291 (CoB6)	oCBM4996	GACCATGCGTTGTTCAAGCG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 247935 in pRALTA_0291 (CoB6)	oCBM4375	GCAATAGAAGTCGGCGGTCT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 259792 in up66-RSp0211 (CoB6)	oCBM4700	GGTTCATCTCGTCCCAGCTC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 259792 in up66-RSp0211 (CoB6)	oCBM4701	CGGACAGATCGGTGTAGACG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 583937 in rhlE1 (CoB7)	oCBM4926	CCAACGTGCCGGAAGACTAT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 583937 in rhlE1 (CoB7)	oCBM4927	TTGAACGGCTGCTTGGGTG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 759825 in nusB (CoB7)	oCBM4708	GTGAAAGGCGCGGATATTGG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 759825 in nusB (CoB7)	oCBM4709	CCTGTACCAATGGCTGCTCA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 267735 in RSc0243 (CoB10)	oCBM4906	GGTGTCCCATGCCGGTTC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 267735 in RSc0243 (CoB10)	oCBM4907	GGAAGGGTTCGACGTGGTG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1848419 in up8-RSp1469 (CoB10)	oCBM4845	TTGCTCATCAACTGGCTCAC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1848419 in up8-RSp1469 (CoB10)	oCBM4846	GCGAAGCGACAGATATGGAT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 197522 in pRalta 0230 (CoB10)	oCBM4904	CATCGAGCCGTATCCGTTCA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 197522 in pRalta_0230 (CoB10)	oCBM4905	GCCCGAAAATACTTGACGCC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 463545 in pRALTA_0552 (CoB10)	oCBM4961	CGACGCACCTTCCTCAAGTA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 463545 in pRALTA_0552 (CoB10)	oCBM4962	TGTAGAACATCTTCGGCCGG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 502781 in pRalta 0615 (CoB10)	oCBM4807	ATCCCTGCATCGTAGAGGTG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 502781 in pRalta_0615 (CoB10)	oCBM4808	GTCACAATCAGCCGGAAACT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1414257 in RSp1116 (CoB10)	oCBM4908	CGCACTATCAGTCCTAGGCC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1414257 in RSp1116 (CoB10)	oCBM4909	CTCGTCGTTGTTGATGTGGC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2141969 in RSc1962 (CoG2)	oCBM4997	TGCATTCAATCCTAGCGCGT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2141969 in RSc1962 (CoG2)	oCBM4661	CGTCACCGGCCAGATCAGTTCC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2157029 in Rsc1982 (CoG2)	oCBM4702	GCCACTTCTTGCCCTCGAA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2157029 in Rsc1982 (CoG2)	oCBM4703	GGCATCGATCTCTGACGCAA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1222748 in ppiB (CoG4)	oCBM4998	GCTTTCGATCACCACGTCTT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1222748 in ppiB (CoG4)	oCBM4090	TTTCCGCCAAGGAACTCAC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2474088 in RSc2277 (CoG4)	oCBM4999	CAATGTCGGCAAGCACTACA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2474088 in RSc2277 (CoG4)	oCBM4080	TCGACGCTCTCGCTCTTCTTGG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1808334 in cysE1 (CoG7)	oCBM4841	GCGGCATGTAGCTCTGGTAG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1808334 in cysE1 (CoG7)	oCBM4842	ATCGGGAAGGATTCCATCAT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 558407 in up85-gloA (CoG7)	oCBM4718	ACTTGTATTCGGGGTTGTCG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 558407 in up85-gloA (CoG7)	oCBM4719	GCTTACCTGCGCTACTCGTC	This study

Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 777392 in up180-RSp0637 (CoG7)	oCBM4698	TCGCAGAACCATGAATCGCA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 777392 in up180-RSp0637 (CoG7)	oCBM4699	CTGGCTGCACTTCGAGGAA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2473174 in RSc2277 (CoG8)	oCBM4839	CGGTCCGGATAGAGTCAAGT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2473174 in RSc2277 (CoG8)	oCBM4840	GTCCTTGAGCTGCAGGATTC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2526677 in RSc2326 (CoG11)	oCBM4710	ACAAGGTGGCAAGTACGTCC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2526677 in RSc2326 (CoG11)	oCBM4711	CCGACGTGAAATCCAGCAAC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1754626 in phbB (CoG13)	oCBM4712	CCCGTGAGTTTTGCGTGAAG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1754626 in phbB (CoG13)	oCBM4713	GGTGGAGTAGTTGGTCTGGC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 215421 in RSc0194 (CoG13)	oCBM4928	GTCTCTCAAGCCCGGTGATC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 215421 in RSc0194 (CoG13)	oCBM4929	ACAGTTCGAATGTGGCAGGT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 931539 in RSp0741 (CoG13)	oCBM4910	GCTCACGTTGGCTCTTGTTG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 931539 in RSp0741 (CoG13)	oCBM4911	GCACTCAAGGATCCGACCAG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1395827 in dnaB (CoG15)	oCBM3654	TCTAAGCACGTACGCAGGAC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1395827 in dnaB (CoG15)	oCBM3655	CCAATGCGTTCAGGTAGGC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1781273 in RSp0208 (CoG15)	oCBM4765	GTTTCGATACCCACGTCGAG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1781273 in RSp0208 (CoG15)	oCBM4766	GATTTTCCATTCGAGCCTCA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1781273 in RSp1422 (CoG15)	oCBM4769	TGAGCATCGAAAAGTGATCG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 1781273 in RSp1422 (CoG15)	oCBM4770	CACGGACCCTACCAGTTGAT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2589097 in RSc2383 (CoG15)	oCBM4837	GTCCTTGGGCAGATTGAACA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 2589097 in RSc2383 (CoG15)	oCBM4838	ACGCTGACGGTGGACTTG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 345083 in pRALTA_0398 (CoG15)	oCBM4696	TGAGAACTTGGACGCGTTGA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 345083 in pRALTA_0398 (CoG15)	oCBM4697	CGCTCCAATCCTTCGACCTT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 579465 in fadE (CoG15)	oCBM3658	GTCCATGGCATCGTTGATCA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 579465 in fadE (CoG15)	oCBM3659	CATCCGGGCGTCAACATC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 713879 in Rsp0574 (CoG15)	oCBM4724	TGCGGTAGTCATCTTGCTTG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 713879 in Rsp0574 (CoG15)	oCBM4725	TCACTGCGACTGAGTTCGAC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 719867 in RSc0672 (CoG13)	oCBM4809	CATCGAACAGGTGATGAACG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 719867 in RSc0672 (CoG13)	oCBM4810	GATGATCTTGATGCCGTCGT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 722803 in RSp0587 (CoG15)	oCBM3620	TGCATCTTCCCGTCCTTGAT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 722803 in RSp0587 (CoG15)	oCBM3621	GTTCGGTGCGGGTTGTATC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 957425 in RSp0759 (CoG15)	oCBM3612	ACATGCTCTTCTCCACCGAA	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 957425 in RSp0759 (CoG15)	oCBM3613	AACTGATCCTGCGTGACCTT	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 973212 inRSp0769 (CoG15)	oCBM3608	GTGTCCCACGTTTGACCATC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 973212 inRSp0769 (CoG15)	oCBM3609	ACCTTCGTTAACTCTGGGGC	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 980536 in up116-RSp0780 (CoG15)	oCBM3604	GGCGGTCGATTTCAGTGAAG	This study
Sequence-based ragent- PCR primer	Amplification of a fragment for sequencing around the mutation position 980536 in up116-RSp0780 (CoG15)	oCBM3605	TCGCCAATGACATCCGGATA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0595 (CoB1, sequence position 728875)	oCBM4343	GGACTCACTCGTAATACGGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0595 (CoB1, sequence position 728875)	oCBM4345	GTTGCCTCTTGACCGTATCG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0595 (CoB1, sequence position 728875)	oCBM4344	GGACTCACTCGTAATACGGT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0595 (CoB1, sequence position 728875)	oCBM4345	GTTGCCTCTTGACCGTATCG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc1598 (CoB2, sequence position 171563)	oCBM4913	CGCGCGCATCGATCAGGA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc1598 (CoB2, sequence position 171563)	oCBM4914	CTGACCGACCGCTACCGCTT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc1598 (CoB2, sequence position 171563)	oCBM4912	CGCGCGCATCGATCAGGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc1598 (CoB2, sequence position 171563)	oCBM4914	CTGACCGACCGCTACCGCTT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0449 (CoB3, sequence position 387383)	oCBM4364	ACCAACATTTTGGCCTCAAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0449 (CoB3, sequence position 387383)	oCBM4365	CCCAAAGACTTACTCAGACCG	This study

Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0788 (CoB3, sequence position 992273)	oCBM4349	CCTCGACATCGTCAATGGGA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0788 (CoB3, sequence position 992273)	oCBM4350	ATCGCGCCGTTCTTCAGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1417 (CoB3, sequence position 1781273)	oCBM4354	ACCCAGATTTCCCCAGATGA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1417 (CoB3, sequence position 1781273)	oCBM4355	GCGCCAAACATCCCATGATC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1590 (CoB3, sequence position 1994246)	oCBM4359	GACGGCGTCGCGCTCGCAAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1590 (CoB3, sequence position 1994246)	oCBM4360	AGAACCCGGCGGCGAAGTAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0449 (CoB3, sequence position 387383)	oCBM4363	ACCAACATTTTGGCCTCAAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0449 (CoB3, sequence position 387383)	oCBM4365	CCCAAAGACTTACTCAGACCG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0788 (CoB3, sequence position 992273)	oCBM4348	CCTCGACATCGTCAATGGGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0788 (CoB3, sequence position 992273)	oCBM4350	ATCGCGCCGTTCTTCAGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1590 (CoB3, sequence position 1994246)	oCBM4358	GACGGCGTCGCGCTCGCAAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1590 (CoB3, sequence position 1994246)	oCBM4360	AGAACCCGGCGGCGAAGTAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1417 (CoB3, sequence position 1781273)	oCBM4353	ACCCAGATTTCCCCAGATGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1417 (CoB3, sequence position 1781273)	oCBM4355	GCGCCAAACATCCCATGATC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0288 (CoB6, sequence position 243268)	oCBM4369	AATGGCCTCGAAAAGATGCA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0288 (CoB6, sequence position 243268)	oCBM4370	CGGGATGCTAGCAATTCGAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0291 (CoB6, sequence position 247935)	oCBM4374	CACTGCTGTTTCCGTTTGGT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0291 (CoB6, sequence position 247935)	oCBM4375	GCAATAGAAGTCGGCGGTCT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1071 (CoB6, sequence position 1353434)	oCBM4379	TCTTCAATGAGAAGGAGATAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1071 (CoB6, sequence position 1353434)	oCBM4380	CAGGCTGGTGATGAGATTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up115-RSc0965 (CoB6, sequence position 1013285)	oCBM2532	TTATGGCGGGCCGGGCGGTA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up115-RSc0965 (CoB6, sequence position 1013285)	oCBM2533	GCAGGTCGAGCAAGCAAAGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up66-RSp0211 (CoB6, sequence position 259792)	oCBM4651	CAGTCTAGGCAAGCGTGCGT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up66-RSp0211 (CoB6, sequence position 259792)	oCBM4652	AGGTGATCGGCGGACAGATG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0288 (CoB6, sequence position 243268)	oCBM4368	AATGGCCTCGAAAAGATGCT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0288 (CoB6, sequence position 243268)	oCBM4370	CGGGATGCTAGCAATTCGAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0291 (CoB6, sequence position 247935)	oCBM4373	CACTGCTGTTTCCGTTTGGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0291 (CoB6, sequence position 247935)	oCBM4375	GCAATAGAAGTCGGCGGTCT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1071 (CoB6, sequence position 1353434)	oCBM4378	TCTTCAATGAGAAGGAGATAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1071 (CoB6, sequence position 1353434)	oCBM4380	CAGGCTGGTGATGAGATTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up115-RSc0965 (CoB6, sequence position 1013285)	oCBM2531	TTATGGCGGGCCGGGCGGTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up115-RSc0965 (CoB6, sequence position 1013285)	oCBM2533	GCAGGTCGAGCAAGCAAAGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up66-RSp0211 (CoB6, sequence position 259792)	oCBM4650	CAGTCTAGGCAAGCGTGCGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up66-RSp0211 (CoB6, sequence position 259792)	oCBM4652	AGGTGATCGGCGGACAGATG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of nusB (CoB7, sequence position 759825)	oCBM4394	CGCGCCCACTTCGATGCCT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of nusB (CoB7, sequence position 759825)	oCBM4395	GGTTCGCCGATGCCCATGTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of rhlE1 (CoB7, sequence position 583937)	oCBM4389	GGCAGTTCGCGCTTGATCTA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of rhlE1 (CoB7, sequence position 583937)	oCBM4390	TGGGCTTCATCCACGACATC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of nusB (CoB7, sequence position 759825)	oCBM4393	CGCGCCCACTTCGATGCCC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of nusB (CoB7, sequence position 759825)	oCBM4395	GGTTCGCCGATGCCCATGTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of rhlE1 (CoB7, sequence position 583937)	oCBM4388	GGCAGTTCGCGCTTGATCTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of rhlE1 (CoB7, sequence position 583937)	oCBM4390	TGGGCTTCATCCACGACATC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0230 (CoB10, sequence position 197522)	oCBM4417	CAGCGTTGGGCTTCCGCCA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0230 (CoB10, sequence position 197522)	oCBM4418	CGACAAAGGACAGCAGCGCG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0552 (CoB10, sequence position 463545)	oCBM4440	AGGTGGGCATGAGTTGCAAA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0552 (CoB10, sequence position 463545)	oCBM4441	CCGAACTGGTGGATGTTGGT	This study

Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0615 (CoB10, sequence position 502781)	oCBM4445	ATTGCGCCCCGTCGGTCACA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA 0615 (CoB10, sequence position 502781)	oCBM4446	CCAGCCAACTACCCGTCGAGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc0243 (CoB10, sequence position 267735)	oCBM3910	CGAGGTGGCCGCCGAACAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc0243 (CoB10, sequence position 267735)	oCBM3911	ATCCGGTTGAGGGGCTTGGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1116 (CoB10, sequence position 1414257)	oCBM4657	ACCGGCATTGGCGGCAAAA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1116 (CoB10, sequence position 1414257)	oCBM4658	GCAGTTCAAGCAGACCGTGGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up8-RSp1469 (CoB10, sequence position 1848419)	oCBM3931	TTTGTCGTTCATGCCGGTAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up8-RSp1469 (CoB10, sequence position 1848419)	oCBM3932	CCTTTCGTTTTCACCGGCAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0230 (CoB10, sequence position 197522)	oCBM4416	CAGCGTTGGGCTTCCGCCG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA 0230 (CoB10, sequence position 197522)	oCBM4418	CGACAAAGGACAGCAGCGCG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0552 (CoB10, sequence position 463545)	oCBM4439	AGGTGGGCATGAGTTGCAAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0552 (CoB10, sequence position 463545)	oCBM4441	CCGAACTGGTGGATGTTGGT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0615 (CoB10, sequence position 502781)	oCBM4444	ATTGCGCCCCGTCGGTCACG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA 0615 (CoB10, sequence position 502781)	oCBM4446	CCAGCCAACTACCCGTCGAGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc0243 (CoB10, sequence position 267735)	oCBM3909	CGAGGTGGCCGCCGAACAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc0243 (CoB10, sequence position 267735)	oCBM3911	ATCCGGTTGAGGGGCTTGGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1116 (CoB10, sequence position 1414257)	oCBM4656	ACCGGCATTGGCGGCAAGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1116 (CoB10, sequence position 1414257)	oCBM4658	GCAGTTCAAGCAGACCGTGGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up8-RSp1469 (CoB10, sequence position 1848419)	oCBM3930	TTTGTCGTTCATGCCGGTAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up8-RSp1469 (CoB10, sequence position 1848419)	oCBM3932	CCTTTCGTTTTCACCGGCAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc1097 (CoG1, sequence position 1155501)	oCBM4455	ATCGCCAACCCGTCGGTGA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc1097 (CoG1, sequence position 1155501)	oCBM4456	GAAACACGTGCGGCTCGGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc1097 (CoG1, sequence position 1155501)	oCBM4454	ATCGCCAACCCGTCGGTGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc1097 (CoG1, sequence position 1155501)	oCBM4456	GAAACACGTGCGGCTCGGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc1962 (CoG2, sequence position 2141969)	oCBM4659	AAGCCGCGCGGCCCGCATG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc1962 (CoG2, sequence position 2141969)	oCBM4661	CGTCACCGGCCAGATCAGTTCC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc1982 (CoG2, sequence position 2157029)	oCBM4337	CCTGTCGATGATCAACGAGAAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc1982 (CoG2, sequence position 2157029)	oCBM4336	TGCGGGACATGGATTGAAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc1962 (CoG2, sequence position 2141969)	oCBM4660	AAGCCGCGCGGCCCGCATA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc1962 (CoG2, sequence position 2141969)	oCBM4661	CGTCACCGGCCAGATCAGTTCC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc1982 (CoG2, sequence position 21570299)	oCBM4338	CCTGTCGATGATCAACGAGAAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc1982 (CoG2, sequence position 21570299)	oCBM4336	TGCGGGACATGGATTGAAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc2277 (CoG4, sequence position 2474088)	oCBM4339	CAACGAATACGGGCTGGTGCT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc2277 (CoG4, sequence position 2474088)	oCBM4080	TCGACGCTCTCGCTCTTCTTGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc2277 (CoG4, sequence position 2474088)	oCBM4340	CAACGAATACGGGCTGGTGCG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc2277 (CoG4, sequence position 2474088)	oCBM4080	TCGACGCTCTCGCTCTTCTTGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of ppiB (CoG4, sequence position 1222748)	oCBM4089	CGACCACGTTGATGAAGAACAA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of ppiB (CoG4, sequence position 1222748)	oCBM4090	TTTCCGCCAAGGAACTCAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of ppiB (CoG4, sequence position 1222748)	oCBM4088	CGACCACGTTGATGAAGAACAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of ppiB (CoG4, sequence position 1222748)	oCBM4090	TTTCCGCCAAGGAACTCAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of cysE1 (CoG7, sequence position 1808334)	oCBM4475	GGGCTGGAACTCGGCATCCA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of cysE1 (CoG7, sequence position 1808334)	oCBM4476	TGTCGGACATGCGGGCGAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0637/RSp0638 (CoG7, sequence position 777392)	oCBM4490	CCACCGCCCTCGACAGGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0637/RSp0638 (CoG7, sequence position 777392)	oCBM4491	TGCACCACTCGAGCGCTTCT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up85-gloA/RSc0521 (CoG7, sequence position 558407)	oCBM4480	GTGGGGGGCGGTTGGTTTAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up85-gloA/RSc0521 (CoG7, sequence position 558407)	oCBM4481	CACGCTTCTGGGACACGGTG	This study

Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of cysE1 (CoG7, sequence position 1808334)	oCBM4474	GGGCTGGAACTCGGCATCCG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of cysE1 (CoG7, sequence position 1808334)	oCBM4476	TGTCGGACATGCGGGCGAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0637/RSp0638 (CoG7, sequence position 777392)	oCBM4489	CCACCGCCCTCGACAGGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0637/RSp0638 (CoG7, sequence position 777392)	oCBM4491	TGCACCACTCGAGCGCTTCT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up85-gloA/RSc0521 (CoG7, sequence position 558407)	oCBM4479	GTGGGGGGCGGTTGGTTTAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up85-gloA/RSc0521 (CoG7, sequence position 558407)	oCBM4481	CACGCTTCTGGGACACGGTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc2277 (CoG8, sequence position 2473174)	oCBM4513	TCCTGGTCAGCGCCTTCTAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc2277 (CoG8, sequence position 2473174)	oCBM4514	TCGGTGGTGATCTGCTCGAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc2277 (CoG8, sequence position 2473174)	oCBM4512	TCCTGGTCAGCGCCTTCTAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc2277 (CoG8, sequence position 2473174)	oCBM4514	TCGGTGGTGATCTGCTCGAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc2326 (CoG11, sequence position2526677)	oCBM4314	GCTTCCAGCACCTGCAGGGT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc2326 (CoG11, sequence position2526677)	oCBM4315	CCACGGTCCATAAGTTCGCGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc2326 (CoG11, sequence position 2526677)	oCBM4313	GCTTCCAGCACCTGCAGGGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc2326 (CoG11, sequence position 2526677)	oCBM4315	CCACGGTCCATAAGTTCGCGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of phbB (CoG13, sequence position 1754626)	oCBM4302	GCCATTTGTCCTTGCGCGGTA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of phbB (CoG13, sequence position 1754626)	oCBM4303	TTGCCATCGGTCACCCCATC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc0194 (CoG13, sequence position 2154217)	oCBM4311	TTGTCGGCGCGGCATGTATAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc0194 (CoG13, sequence position 2154217)	oCBM4312	CAGTCGACGCCCATGCCATC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc0672 (CoG13, sequence position 719867)	oCBM4669	CACAGGCTCGACAGGCGGA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc0672 (CoG13, sequence position 719867)	oCBM4670	CGATCTGGACCTGCTGACGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of Rsp0741/tssJ (CoG13, sequence position 931539)	oCBM4305	TCGCGCAAGCGTGCAATCCCT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of Rsp0741/tssJ (CoG13, sequence position 931539)	oCBM4306	CCAACCTGAACGCGGCGAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of phbB (CoG13, sequence position 1754626)	oCBM4301	GCCATTTGTCCTTGCGCGGTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of phbB (CoG13, sequence position 1754626)	oCBM4303	TTGCCATCGGTCACCCCATC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc0194 (CoG13, sequence position 215421)	oCBM4310	TTGTCGGCGCGGCATGTATAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc0194 (CoG13, sequence position 215421)	oCBM4312	CAGTCGACGCCCATGCCATC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc0672 (CoG13, sequence position 719867)	oCBM4668	CACAGGCTCGACAGGCGGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc0672 (CoG13, sequence position 719867)	oCBM4670	CGATCTGGACCTGCTGACGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0741/tssJ (CoG13, sequence position 931539)	oCBM4304	TCGCGCAAGCGTGCAATCCCA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0741/tssJ (CoG13, sequence position 931539)	oCBM4306	CCAACCTGAACGCGGCGAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of dnaB (CoG15, sequence position 1395827)	oCBM4269	CCGATCCCCAGCTCGAATGT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of dnaB (CoG15, sequence position 1395827)	oCBM4270	ATGCGGTCCAGGTCAACGAA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of fadE (CoG15, sequence position 579465)	oCBM4272	TGGACGACGGCAGCGAGAAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of fadE (CoG15, sequence position 579465)	oCBM4273	GGTGATGAAGCTGGCGACGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0398 (CoG15, sequence position 345083)	oCBM3952	TGACGCCGTGCGTGAATAAA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of pRALTA_0398 (CoG15, sequence position 345083)	oCBM3953	CCTTGGAGATGCTCGCACTC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc2383 (CoG15, sequence position 2589097)	oCBM4275	CCCTGGTCGTACAGCGCGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSc2383 (CoG15, sequence position 2589097)	oCBM4276	TCGAATCGGACCAGGGCCAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0208 (CoG15, sequence position 257773)	oCBM4287	CCATGATCCCCGGCAAAATA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0208 (CoG15, sequence position 257773)	oCBM4288	CAAGGCAAAGTCCCAACAGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0574 (CoG15, sequence position 713879)	oCBM4278	ATTTGCATGTGATGGCGGAA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0574 (CoG15, sequence position 713879)	oCBM4279	CAACACCATCAAACGCTCGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0587 (CoG15, sequence position 722803)	oCBM4284	GGAACGCAGCGCGGGATAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0587 (CoG15, sequence position 722803)	oCBM4285	CCTGCCCCACCTAATGCGAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0759 (CoG15, sequence position 957425)	oCBM3752	AGTATTGGGAGCACCTCTTTT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0759 (CoG15, sequence position 957425)	oCBM3753	GGGCTTGGTTTCCGTCTTG	This study

Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0769 (CoG15, sequence position 973212)	oCBM3762	TTTCACCGTGACTCTGGTAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp0769 (CoG15, sequence position 973212)	oCBM3763	GGATGCCAGTTAGCCGATTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1422 (CoG15, sequence position 1789297)	oCBM3782	ATCGCGCCGAGCAACCAAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of RSp1422 (CoG15, sequence position 1789297)	oCBM3783	CTGTACGACCACCCGACG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up116-RSp0780 (CoG15, sequence position 980536)	oCBM3767	CGAATGGCCAAGAACATGGT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the mutant allele of up116-RSp0780 (CoG15, sequence position 980536)	oCBM3768	AGGTGGCCTTGTTCAAATCG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of dnaB (CoG15, sequence position 1395827)	oCBM4268	CCGATCCCCAGCTCGAATGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of dnaB (CoG15, sequence position 1395827)	oCBM4270	ATGCGGTCCAGGTCAACGAA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of fadE (CoG15, sequence position 579465)	oCBM4271	TGGACGACGGCAGCGAGAAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of fadE (CoG15, sequence position 579465)	oCBM4273	GGTGATGAAGCTGGCGACGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0398 (CoG15, sequence position 345083)	oCBM3951	TGACGCCGTGCGTGAATAAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of pRALTA_0398 (CoG15, sequence position 345083)	oCBM3953	CCTTGGAGATGCTCGCACTC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc2383 (CoG15, sequence position 2589097)	oCBM4274	CCCTGGTCGTACAGCGCGA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSc2383 (CoG15, sequence position 2589097)	oCBM4276	TCGAATCGGACCAGGGCCAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0208 (CoG15, sequence position 257773)	oCBM4286	CCATGATCCCCGGCAAAATG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0208 (CoG15, sequence position 257773)	oCBM4288	CAAGGCAAAGTCCCAACAGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0574 (CoG15, sequence position 713879)	oCBM4277	ATTTGCATGTGATGGCGGAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0574 (CoG15, sequence position 713879)	oCBM4279	CAACACCATCAAACGCTCGG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0587 (CoG15, sequence position 722803)	oCBM4283	GGAACGCAGCGCGGGATAA	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0587 (CoG15, sequence position 722803)	oCBM4285	CCTGCCCCACCTAATGCGAG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0759 (CoG15, sequence position 957425)	oCBM3751	AGTATTGGGAGCACCTCTTCC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0759 (CoG15, sequence position 957425)	oCBM3753	GGGCTTGGTTTCCGTCTTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0769 (CoG15, sequence position 973212)	oCBM3761	TTTCACCGTGACTCTGGTAT	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp0769 (CoG15, sequence position 973212)	oCBM3763	GGATGCCAGTTAGCCGATTG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1422 (CoG15, sequence position 1789297)	oCBM3781	ATCGCGCCGAGCAACCAAC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of RSp1422 (CoG15, sequence position 1789297)	oCBM3783	CTGTACGACCACCCGACG	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up116-RSp0780 (CoG15, sequence position980536)	oCBM3766	CGAATGGCCAAGAACATGGC	This study
Sequence-based ragent- PCR primer	Screening oligos for detecting the wild-type allele of up116-RSp0780 (CoG15, sequence position980536)	oCBM3768	AGGTGGCCTTGTTCAAATCG	This study
Software, algorithm	breseq	RRID:SCR_010810	v0.33.1	PMID: 24838886
Software, algorithm	IGV	RRID:SCR_011793	v2.5.2	PMID: 21221095
Software, algorithm	R project for statistical computing	RRID:SCR_001905	v3.6.1	R Core Team
Software, algorithm	R package ggplot2	RRID:SCR_014601	v3.3.5	https://ggplot2.tidyverse.org
Software, algorithm	R package RColorBrewer	RRID:SCR_016697	v1.1-2	https://cran.r-project.org/web/packages/R
Software, algorithm	R package MullerPlot		v0.1.2	https://CRAN.R-project.org/package=MullerP
Software, algorithm	R package dplyr	RRID:SCR_016708	v0.8.3	https://CRAN.R-project.org/package=dplyr
Software, algorithm	R package tidyverse	RRID:SCR_019186	v1.3.0	https://doi.org/10.21105/joss.01686org
Software, algorithm	R package data.table		v1.12.8	https://CRAN.R-project.org/package=data.ta
Software, algorithm	R package wesanderson		v0.3.6	https://CRAN.R-project.org/package=wesand
Software, algorithm	R package mvtnorm		v1.1-1	http://CRAN.R-project.org/package=mvtnor
Software, algorithm	R package rlist		v0.4.6.1	https://CRAN.R-project.org/package=rlist
Software, algorithm	R package stringr		v1.4.0	https://CRAN.R-project.org/package=stringr

TriR, trimethoprim resistance. KanR, kanamycin resistance. StrR, streptomycin resistance. IGglmS: insertion in the intergenic region downstream the *glmS* gene.

S2.7 Table. Parameter values used in computer simulations

Variable parameters		
Parameter	Definition	Values[a]
B	Number of nodules (size of nodulation bottleneck)	10, 30, 100, **300**, 1000 or 3000
$\omega_n 0$	Nodulation competitiveness of the ancestor	**10^{-4}**, 10^{-3}, or 10^{-2}
$\omega_p 0$	Within-nodule proliferation fitness of the ancestor	**10^{-4}**, 10^{-3}, or 10^{-2}
k	Dimension of the Nodulation competitiveness phenotype	3, 6, 8, **10** or 20
l	Dimenion of the Proliferation phenotype	3, 6, 8, **10** or 20
m	Dimension of the pleiotropic Nodulation competitiveness/Proliferation phenotype	**0**, 2 or 4
Fixed parameters		
Parameter	Definition	Value
N_{rhizo}	Number of bacteria in the rhizosphere	10^{+6}
N_{nod}	Maximum number of bacteria per nodule	10^{+8}
μ_{rhizo}	Mutation rate in the rhizosphere	10^{-2}
N_{cycles}	Number of evolutionary cycles	50
α_{fit}	Shape of inversed Gamma distribution of fitness effects	2.5
β_{fit}	Scale of inversed Gamma distribution of fitness effects	0.15
α_{mut}	Shape of inversed Gamma distribution of mutation effects	2.04
β_{mut}	Scalee of inversed Gamma distribution of mutation effects	0.0208

[a]Bold values mark the default values used in most simulations

Supplementary File 1: Model framework and supplementary simulations to assess the robustness to variations in the genetic architecture of symbiotic traits

August 8, 2022

1 Model framework

1.1 Ecology of symbiotic life cycles

We simulated bacterial populations that repeatedly cycle between two compartments, the rhizosphere and root nodules (within-host compartment; S5 Fig). In the rhizosphere, bacteria undergo transient hypermutagenesis, a phenomenon increasing basal mutation rate by 10-100 times [21]. Although we previously estimated experimentally that bacteria undergo a few (< 5) divisions in the rhizosphere, we were unable to detect any sign of selection associated with these divisions. For simplicity, we thus assume in the simulations that bacteria do not reproduce in this compartment. Bacterial population size in the rhizosphere was set to $N_{rhizo} = 10^6$ cells. To induce and enter nodules, bacteria undergo a very stringent bottleneck. Indeed, we previously determined experimentally that 97% of *M. pudica* nodules contain only one single genotype when co-inoculated with 2 isogenic strains carrying different fluorescent reporters [33], and the number of nodules collected at each cycle of our evolution experiment was generally comprised between 100 and 225 (first and third quartile of the distribution of nodule numbers collected during the evolution experiment, S1 Table), since each plant forms multiple nodules and the experiment was run on pools of 20 or 30 plants. In the model, we thus considered that each nodule is founded by one single bacterium, and we studied the influence of the size of the nodulation bottleneck on adaptation by varying the total number of nodules formed by each population at each cycle (with nodule numbers ranging from 10 to 3000). Moreover, we also know from previous work that bacterial genotypes can significantly differ in their competitive ability to form nodules [17]. Therefore, nodulation represents a selective bottleneck and we define nodulation competitiveness fitness value as the probability that a rhizospheric bacterium forms a nodule. Of note, the size of the nodulation bottleneck is a constant value set within each simulation run, and does not depend on nodulation competitiveness of the population, that evolves along successive cycles. In the model, each nodule is founded by a bacterial genotype sampled from the whole population, with a probability of each genotype forming a nodule equals to the product of its relative frequency in the population by its nodulation competitiveness fitness value. Once in the nodule, bacteria multiply clonally (mutation rate within nodules is negligible compared to that of the rhizosphere [21]. We previously showed that within host fitness values are also dependent on bacterial genotypes (reviewed in [99]), so the number of each bacterial clone within nodule is set to their within-host proliferation fitness value multiplied by the maximum bacterial load per nodule, estimated to $N_{nod} = 10^8$ cells/nodule [33]. At the end of each cycle, bacteria are pooled from all the nodules within a given population and 10^6 cells are randomly sampled to re-inoculate the next cycle. This main version of the model is named "v1".

A variation from the main model was designed to test if the chronology of symbiotic events affects evolutionary dynamics. In this version of the model (named version "Inv1"), we simulate populations that multiply clonally (up to numbers defined by the within-host proliferation fitness value) before experiencing the nodulation bottleneck.

1.2 Evolution of bacterial phenotypes

Using the model from Tenaillon *et al.* as a base [100], we developed a multi-dimensional Fisher's Geometric Model to simulate the adaptive walk of bacteria. We decomposed bacterial fitness into two components (nodulation competitiveness and proliferation within nodules), each described by a phenotypic vector (x and y, respectively):

$$x = \{x_1, ..., x_k, z_1, ..., z_m\}$$

$$y = \{y_1, ..., y_l, z_1, ..., z_m\}$$

where k is the number of dimensions of nodulation-specific phenotypes, l is the number of dimensions of proliferation-specific phenotypes, and m is the number of dimensions of pleiotropic phenotypes (e.g. affecting both nodulation competitiveness and proliferation). These pleiotropic phenotypes can be used to simulate possible genetic couplings between the two components.

The fitness of a bacterial strain is calculated as $w = w_n \times w_p$ with

$$\begin{cases} w_n = & \exp(-\frac{1}{2}x^T \mathbf{V}_n^{-1} x) \\ w_p = & \exp(-\frac{1}{2}y^T \mathbf{V}_p^{-1} y) \end{cases}$$

the nodulation and proliferation fitness, respectively. Fitness is therefore maximal ($w = 1$) when all traits are zero. $\mathbf{V}_n$ and $\mathbf{V}_p$ are diagonal matrices whose diagonal elements, $\sigma^2_{n_i}$ and $\sigma^2_{p_i}$, determine how deviations from zero for a given trait impact w_n and w_p: the larger these elements, the least the trait influences fitness. For each simulation, the $\sigma^2_{n_i}$ ($\sigma^2_{p_i}$ respectively) elements are randomly drawn from an inverse gamma distribution inv-$\Gamma(\alpha_{fit}, \beta_{fit})$ which average and variance are given by $\beta_{fit}/(\alpha_{fit} - 1)$ and $\beta_{fit}/((\alpha_{fit} - 1)^2(\alpha_{fit} - 2))$, respectively.

Each single mutation will modify all phenotypic components (i.e. all values of vectors x and y) by random values drawn from $k + l + m$ independent centered Gaussian distributions. For each simulation, the variances of these Gaussian distributions are independently drawn from an inverse gamma distribution with parameters α_{mut} and β_{mut}.

1.3 Parameter values

The full list of parameter values used in the simulations is shown in S7 Table. Simulations were used to explore the effect of the size of the nodulation bottleneck and of the fitness of the ancestral strain on bacterial adaptation. We used nodulation bottlenecks of sizes B = 10, 30, 100, 300, 1000, or 3000 nodules. The combinations of ancestral fitness values of nodulation competitiveness and within-nodule proliferation (w_n0 / w_p0) tested were (10^{-4}/ 10^{-4}), (10^{-3}/ 10^{-4}), (10^{-2}/ 10^{-4}),(10^{-4}/ 10^{-3}) and (10^{-4}/ 10^{-2}). We used ancestral fitness values of (10^{-4}/ 10^{-4}) when testing the effect of bottleneck size, and a bottleneck of 300 nodules when testing the effect of ancestral fitness values. The parameters that remained constant in all simulations were either set to match (or approach) experimental estimations (number of bacteria in the rhizosphere $N_{rhizo} = 10^6$, maximum number of bacteria per nodule $N_{nod} = 10^8$, mutation rate $\mu_{rhizo} = 10^{-2}$), or chosen arbitrarily so

that the adaptive walks proceed with a kinetics comparable to the ones observed experimentally ($\alpha_{fit} = 2.5,\ \beta_{fit} = 0.15,\ \alpha_{mut} = 2.04,\ \beta_{mut} = 0.0208$).

In addition, we tested if our hypotheses regarding the genetic architecture of symbiotic traits affected simulation results by modifying the distribution of fitness effects and the level of pleiotropy of simulated mutations. In Fisher's Geometric Model, the distribution of mutational effects is dependent on the number of dimensions of the phenotypic space [101] (parameters k, l, and m in our model, where k is the dimension of nodulation-specific phenotypes, l is the dimension of proliferation-specific phenotypes, and m is the dimension of pleiotropic phenotypes). While we used values of $k = l = 10$ and $m = 0$ in most simulations, the proportion of beneficial and deleterious mutations was altered by setting $k = l = 3$ or $k = l = 20$ (S6 Fig). Partial pleiotropy between the 2 phenotypic components manifests itself by a positive correlation between mutational effects (S7 Fig), which was achieved by using non-null values for m while maintaining a total number of dimensions constant (hence keeping the same proportion of beneficial or deleterious mutations). We used the following combinations of parameters: $k = l = 6$ and $m = 4$, and $k = l = 2$ and $m = 8$.

When analyzing the 'Inv1' version of the model, we studied the effect of the size of the nodulation bottleneck (B = 10-3000); default values were used for all other parameters ($\omega_n 0 = 10^{-4}$, $\omega_p 0 = 10^{-4}$, $k = l = 10$ and $m = 0$).

1.4 Simulations

We performed 100 replicate simulations of 50 evolutionary cycles for each set of parameters tested. Each replicate simulation begins by defining the variance-covariance matrices of mutation and fitness effects and the exact fitness values of the ancestor. The latter was achieved by following two steps: (i) defining random vector values for phenotypic traits (from a normal distribution, $\mu = 0,\ \sigma = 0.02$) and (ii) finding a set of vector values that minimize the sum of squared differences to the target value of each fitness component ('optim' function in R). This procedure allowed us to start each simulation with randomized vectors of phenotypic values while still being able to set manually the values of both fitness components.

All information regarding population size and composition at each cycle was recorded during simulations. To generate figures, we computed the mean fitness values of each replicate population at each cycle, as well as the fitness effect of each mutation on the two phenotypic components. Mean population fitness values are shown either individually (S8-S16 Figs) or as the median value of the 100 replicate simulations (Figure 6 of the main text; S9, S10, S12-S15 Figs). To represent the relative fitness effect of selected mutations on the two phenotypic components, we computed the fold effect of each mutation on nodulation competitiveness over its effect on proliferation.

All simulations and data analyses were performed with R v3.6.1 [102] and the following packages: *mvtnorm* [103], *rlist* [104], *ggplot2* [105], *tydiverse* [106] and *stringr* [107]. R code to perform and analyze these simulations, together with simulations results needed to reproduce the figures, are available from the Data INRAE Dataverse : https://doi.org/10.15454/QYB2S9.

2 Supplementary simulations show that results are robust to variations in the genetic architecture of symbiotic traits

The main objective of these evolutionary simulations was to investigate the trajectory of bacterial adaptation to symbiotic life cycles, and its dependence on selected ecological and evolutionary parameters. In the main text of this manuscript, we focused on the analysis of two parameters: the

size of the nodulation bottleneck and the initial fitness values of the ancestor. Evolutionary simulations relied on a generic model of adaptation (Fisher's Geometric Model, FGM) that has proven to provide a good qualitative description of adaptation when compared to microbial evolution experiments [101]. However, this theoretical model makes assumptions on the genetic architecture of the phenotypic traits of interest (the distribution of fitness effects and the level of pleiotropy) that are very hard to measure experimentally, and we wondered if the choice of values for some of the specific evolutionary parameters used in this model could alter its predictions. Therefore, we wanted to assess the robustness of our results to these assumptions. We ran additional simulations to test the effect of (i) the number of phenotypic dimensions of the FGM (controlling the shape of the distributions of mutational effects, hence the proportion of beneficial, neutral, and deleterious mutations) and (ii) the level of pleiotropy of mutations (no pleiotropy vs. partial pleiotropy).

First, the number of dimensions for nodulation competitiveness and within-host proliferation was set to 3 or 20 (compared to 10 for the figure in the main text). We observed that increasing the dimensions of phenotypic vectors from 3 to 20 decreased the proportion of beneficial mutations from ∼46% to ∼20%, respectively, in the low-fitness ancestor (S6 Fig), and decreased the rate of adaptation (S8-S13 Figs). We still observed an asymmetry between selection for nodulation competitiveness and within-host proliferation, that was most prevalent for smaller nodulation bottlenecks (S8-S10 Figs) or when the initial nodulation competitiveness fitness was lower than the initial proliferation fitness (S11-S13 Figs). These results support the conclusion that the asymmetry between selective forces in this experimental life cycle can be expected for phenotypic traits showing a wide diversity of genetic architectures. We note that a difference in the number of dimensions between nodulation competitiveness and proliferation within nodules could also induce such an asymmetry, since the trait with the lowest dimension would have access to more beneficial mutations and would likely improve faster. Although we cannot rule out that symbiotic traits have different distributions of fitness effects in our evolving lineages, simulation results indicate that a selective nodulation bottleneck (when occurring just after the genetic diversification phase) is a sufficient condition to promote the faster selection of nodulation competitiveness, and thus likely plays a role in explaining our experimental observations.

Another genetic factor that can affect evolutionary trajectories is the level of pleiotropy of mutations, *i.e.* the fact that one given mutation can alter the two traits. Pleiotropy can be qualified either as positive, when a mutation improves the two (or more) traits of interest, or as antagonistic, when a mutation improves one trait and worsens the other. Since our experimental results identified cases of positive pleiotropy (Figures 4 and 5, main text), we ran additional simulations to test its effect on the evolutionary trajectories of bacterial populations. This was done by using non-zero values for the number of dimension of pleiotropy phenotypes ($m > 0$), which altered the distributions of fitness effects by introducing a positive correlation between the two phenotypes (S7a-c Fig). A moderate level of pleiotropy ($k = l = 6$ and $m = 4$) has little impact on adaptive trajectories (S14 Fig). Increasing the level of pleiotropy ($k = l = 2$ and $m = 8$) decreases asymmetry between selection on the two phenotypic traits, and accelerates adaptation (S15 Fig). This was expected since a higher correlation between the two traits leads to increased variance in the distribution of fitness effects (S7d Fig). Therefore, more mutations with high-fitness effects are available for natural selection to act on.

Conclusions and Discussion

FGM provides a simple framework that allows to reproduce some general patterns observed in experimental evolution studies. However, the estimation of evolutionary parameters from experimental data is not straightforward, and it is thus difficult to generate quantitative predictions.

Here we explored different combinations of evolutionary parameters to test if they affected the main results of our simulations. In particular, we modified the proportion of beneficial vs. deleterious mutations available for each phenotypic component, by changing the dimensionality of the model, and the degree of pleiotropy of individual mutations. Overall, these additional simulations qualitatively supported the conclusions presented in the main text, indicating their robustness to a range of assumptions regarding the genetic architectures of the two traits of interest.

Our simulation and experimental results indicate that selective bottlenecks can play a significant role in the evolution of symbiotic micro-organisms. While the literature on the evolution of pathogenicity acknowledges that dissemination can be an evolvable trait [108], the stochasticity associated with strong infection bottlenecks is seldom considered, although it can have important consequences [57].Consider for example a bacterial population where two independent mutants arise, the first one having a 10-fold increase in competitiveness for host entry and the other one a 10-fold increase in within-host proliferation. In theory, these two mutants have the same fitness (product of competitiveness for host entry by within-host proliferation phenotypic values) and should have the same probability to invade the resident population. However, if the selective bottleneck is strong and occurs before proliferation within the host, the "competitiveness mutant" will have a much higher probability to avoid extinction (and thus, a higher "realized fitness") than the "proliferation mutant". On the contrary, when clonal expansion occurs before the bottleneck, these probabilities become equal, since the lower competitiveness of the "proliferation mutant" will be compensated for by the higher number of individuals. This process, explaining the asymmetry between selection for host entry and selection for within-host proliferation, is thus dependent on the chronology of events and will become prominent when adaptive mutants are rare and for strong bottleneck sizes (but disappears for weak bottlenecks due to averaging effects).

CHAPTER 3

Functional convergences in adaptive evolution transforming a plant pathogen into legume symbionts

Abstract

Rhizobia, the nitrogen-fixing symbionts of legumes, likely emerged from horizontal transfers of essential symbiotic genes followed by additional adaptation steps to endosymbiosis under plant selection pressure. To get insights into the biological functions required or deleterious for bacteria to establish symbiosis, we characterized seven highly adaptive mutations that contributed to the progressive conversion of a plant pathogen into intracellular legume symbionts following the acquisition of symbiotic genes. We found that most of the mutations decreased the pathogenicity of the ancestral strain towards its original host plant, showing a strong antagonism between pathogenicity and symbiosis establishment. Moreover, convergent evolution towards an increase in flagellar motility, a decrease in EPS production as well as an enhanced carbon and nitrogen metabolism and tolerance to acidic conditions were observed as a consequence of multiple independent, *a priori* unrelated mutations from 2 independent lineages. The potential role of these bacterial functions in symbiotic adaptation is discussed.

Introduction

During the course of evolution, many bacterial lineages evolved the capacity to interact with eukaryotic hosts. These associations range from pathogenicity to commensalism to mutualism, with different degrees of dependence and dominance between partners, forming a fluid continuum, along which bacteria can move. Movements along this continuum depend on ecological and environmental conditions, including the presence of other microorganisms in addition to the two partners of an interaction. These abiotic and biotic factors may strongly influence the strength and the direction of those positive or negative interactions (Brenner et al., 2021; Flórez et al., 2017; Hosokawa et al., 2016; King et al., 2016). Consequently, the terms “parasite” and “mutualist”, instead of representing two sides of a binary system, represent two extremes of a wide range of possible interactions between bacteria and their hosts and are defined by the net-value of the costs-benefits of the relationship for both partners (Drew et al., 2021). In parasitic interactions, the fitness improvements for one partner represent energetic costs for the other one while in mutualistic interactions gains of fitness are provided for both partners once the symbiosis is established. Besides environmental factors, genetic factors such as the horizontal transfers of virulence or mutualistic traits or the acquisition of adaptive mutations, can determine the direction and the speed of changes in bacterial lifestyle

Table3.1. Symbiotic phenotypes of beneficial mutations (Doin de Moura et al., 2022)

Mutation[a]	Description -gene	Mean symbiotic CI[b]	Mean CI for rhizosphere colonization[c]	Mean CI for nodulation competitiveness[d]	Mean gain factor in within-host proliferation[e]	Improved symbiotic traits[f]
RSp0595-P132P	Putative N-acyl-homoserine lactonase	77.0	1.0	5.2	1.6	Nod+
RSc1598-P455L	Transmembrane sensor histidine kinase	24.2	1.4	4.7	6.4	Nod+ Pro+
up^{115}-RSc0965	Unknown function, negatively regulates *efpR* expression	31.1	2.9	8.2	2.6	Rhizo+ Nod+ Pro+
rhlE1 -L368L	ATP-dependent RNA helicase protein	46.2	4.0	28.2	2.4	Rhizo+ Nod+ Pro+
efpR -E66K	Transcriptional regulator	212.1	4.0	35.0	3.7	Rhizo+ Nod+ Pro+
ppiB -Q112*	Peptyl-prolyl-cis/trans isomerase	24.5	1.0	5.0	1.4	Nod+
RSc2277-V321G	Putative transporter protein	133.9	10.1	58.0	2.5	Nod+ Pro+

[a]up^{x}, intergenic mutations located x nucleotides upstream the gene indicated. [b]Mean values of competitive indexes (CI) obtained from 3 to 4 independent experiments. [c]Mean values obtained from at least 15 measurements from 3 to 4 independent experiments. [d]Mean values of competitive indexes (CI) obtained from 3 to 4 independent experiments. Figures in bold are statistically different from 1 (Student t-test $P<0.05$). [e]Mean values obtained from at least 15 measurements from 3 to 4 independent experiments. Figures in bold are statistically different from 1 (Wilcoxon test, $P<0.05$ with the Benjamini-Hochberg correction). [f]Rhizo+, improvement in rhizosphere colonization. Nod+, improvement in nodulation competitiveness. Pro+, improvement in within-host proliferation.

along the continuum (Brenner et al., 2021; Melnyk et al., 2019). Host-associated lifestyle changes were observed several times in proteobacteria. Strikingly, in this branch, mutualistic bacteria seem frequently derived from parasitic ancestors, several examples of mutualistic taxa being completely nested in parasitic clades in phylogenetic trees (Sachs et al., 2014). Conversely, mutualists exhibit rare reversions to other lifestyles, either parasitic or free-living, suggesting host-controlled stabilization mechanisms of mutualism.

Among beneficial interactions, the symbioses between legumes and nitrogen fixing bacteria have been extensively studied (Oldroyd and Downie, 2008; Poole et al., 2018). These interactions are characterized by complex molecular and morphological changes in both partners, resulting in the formation of a new organ in the root of the plant partner, the nodule. The interaction starts when specific flavonoids exuded by the plant trigger the production of rhizobial signal molecules, the Nod Factors, that induce a curling of root hairs around the attached bacterium. Most of the time, only one bacterium is entrapped within the curled root hair and initiates an infection site. Then, the bacterium divides clonally and progress in root hairs through the formation of tubular structures, called infection threads. These structures cross the different root cell layers, which at the same time divide to form the nodule. In the second part of this interaction, bacteria are released from infection threads in the cytoplasm of nodule cells where they persist and differentiate into nitrogen-fixing bacteroids. All over the interaction process, the plant partner can check the symbiotic status of the bacteria and controls all stages of the establishment of the symbiosis (Daubech et al., 2017; Masson-Boivin and Sachs, 2018; Regus et al., 2014; Westhoek et al., 2021).

Rhizobia, the bacteria able to develop this complex interaction with legume plants, form a phylogenetically diverse group of bacteria belonging to at least 18 different genera among α and β-proteobacteria (Masson-Boivin et al., 2009). With a few exceptions (Giraud et al., 2007), this process occurs thanks to a set of essential bacterial symbiotic genes, the *nod* and *nif-fix* genes, often located on mobile genetic elements (MacLean et al., 2007). Genomic studies indicate that the spread of rhizobia in many different taxa occurred via repeated and independent horizontal gene transfer of these symbiotic genes (Andrews et al., 2018). However, horizontal transfer of symbiotic genes is not always sufficient to directly convert a recipient bacterium into a legume symbiont, especially when the transfer occurs between phylogenetically distant bacteria (Nakatsukasa et al., 2008; Nandasena et al., 2007; Wulandari et al., 2022). In these cases, the evolution of rhizobia likely require further adaptation steps to

activate and optimize the symbiotic potential under plant-mediated selection pressure (Masson-Boivin et al., 2009). To validate this two-step scenario, we previously replayed the evolution of a novel rhizobium genus under laboratory conditions by experimental evolution (Doin de Moura et al., 2020). Two phylogenetically related bacterial species but having two contrasted lifestyles, *Cupriavidus taiwanensis* LMG19424*,* the natural symbiont of *Mimosa pudica* (Amadou et al., 2008; Chen et al., 2001a)*,* and *Ralstonia solanacearum* GMI1000 (Salanoubat et al., 2002)*,* a strictly extracellular broad-host range phytopathogen, were used as donor and recipient of the symbiotic genes, respectively (Marchetti et al., 2010). The GMI1000 strain of *R. solanacearum* harboring the *C. taiwanensis* symbiotic plasmid pRalta, initially not able to nodulate *M. pudica*, was progressively adapted to symbiosis with *M. pudica* through repeated nodulation cycles (Marchetti et al., 2017). The two first steps of symbiosis, *i.e*. the capacity to form nodules and infect these nodules intracellularly, were rapidly obtained and strongly improved in most evolved lineages. However, mutualism was not achieved in this experiment after 35 cycles (Doin de Moura et al., 2022). Previous works have identified the genetic bases of this adaptation. Mutations inactivating the Type Three Secretion System of *R. solanacearum*, either *hrcV** or *hrpG**, conferred the ability to nodulate to the strain GMI1000pRalta (Marchetti et al., 2010) while convergent mutations affecting two regulatory pathways, the EfpR and the PhcA pathways, strongly improved the nodulation and the nodule intracellular infection capacities of evolved bacteria (Capela et al., 2017; Doin de Moura et al., 2022; Tang et al., 2020). EfpR and PhcA are global regulators each controlling positively or negatively hundreds of genes, *ca.* 150 of which are in common (Capela et al., 2017; Khokhani et al., 2017; Mori et al., 2018; Perrier et al., 2018; Perrier et al., 2016). However, the EfpR- and PhcA-regulated functions specifically involved in symbiotic adaptation have not been identified. More recently, an extensive genetic analyses of two lineages allowed the identification of 25 adaptive mutations, which contributed to the conversion of *R. solanacearum* into *Mimosa* symbionts (Doin de Moura et al., 2022). Among them, seven are highly adaptive, as they improved the *in planta* fitness of evolved clones by a factor above 20-fold. Two of these seven mutations, *efpR*-E66K and up^{115}-RSc0965, were previously characterized and both repress the EfpR pathway (Capela et al., 2017). The five others affect uncharacterized or poorly characterized genes (Table 3.1). In order to understand the biological functions underlying adaptation to symbiosis and lifestyle change, we evaluated their effect on the pathogenicity of *R. solanacearum* and initiated their functional

characterization. Our data showed that almost all the most adaptive mutations for symbiosis decrease the pathogenicity of the ancestral strain and are pleiotropic, affecting several convergent biological functions in both lineages.

Material and methods

Bacterial strains and growth conditions

Strains and plasmids used in this study are listed in Sup Table 3.1. *Ralstonia solanacearum* strains were grown at 28°C either on BG medium (Boucher et al., 1985) or on synthetic MP medium (Plener et al., 2010) supplemented with 2% glycerol (for bacterial transformation) or with 10 mM of L-glutamate (for motility and EPS production). *Escherichia coli* strains were grown at 37°C on Luria-Bertani medium. Antibiotics were used at the following concentrations: trimethoprim at 100 µg/ml, spectinomycin at 40 µg/ml, and kanamycin at 25 µg/ml (for *E. coli*) or 50 µg/ml (for *R. solanacearum*).

Construction of *Ralstonia solanacearum* mutants

Adaptive mutations previously identified (Doin de Moura et al., 2022) and constitutively expressed reporter genes (GFP or mCherry) were introduced into the GMI1000pRalta strain using the MuGent technique (Dalia et al., 2014) as previously described (Doin de Moura et al., 2022) and verified by Sanger sequencing. Oligonucleotides used for these constructions are described in (Doin de Moura et al., 2022).

Pathogenicity assays

Pathogenicity assays were performed on susceptible *Arabidopsis thaliana* Col-0 plants, as described by (Deslandes et al., 1998). *A. thaliana* seeds were sterilized for 15 minutes in 2.4% sodium hypochlorite solution, then washed several times with sterile water and let to growth in MS medium (Murashige & Skoog, 1962). Sterilized seed were incubated at 4°C for 48 hours, then at 20°C in culture chamber with 16h/8h (light/dark). After 7 days of growth at 20°C, the plantlets were transferred to peat pellets (Jiffy, Netherlands), and grown at 22°C for 18 days, then transferred at 28°C under a 16h/8h (light/dark) photoperiod and light intensity at 500 µE $s–1\ m^{-2}$ for 36 hours before inoculation. After this time, plants were exposed to the different *R. solanacearum* strains by cutting 1-2 cm of roots from the bottom of the Jiffy pot and incubation for 30 minutes in 500 mL of bacterial suspension (5.10^7 CFU/mL). Disease

symptoms were observed and scored for three weeks after inoculation according to the following disease index scale: healthy plants (0), plants with 25%of the leaves wilted (1), plants with 26 to 50%of the leaves wilted (2), plants with 51 to 75%of the leaves wilted (3), and plants having more than 75% of the leaves wilted (4). Pathogenicity assays were repeated three to four times independently for each bacterial strain, and 16 plants were tested per replicate. For Kaplan–Meier survival analyses, data obtained from these assays were transformed into a binary index: 0 for plants showing less than 75% of the leaves wilted and 1 for plants showing 75% or more of the leaves wilted. Hazard (slope of survival curves) ratios between mutant and parental strains were calculated for each independent experiment and the Gehan-Breslow-Wilcoxon test was applied to compute the *P*-values (Remigi et al., 2011). Statistical analyses were performed with GraphPad Prism version 5.00.

Motility assays

Methods for motility assays were adapted from (Perrier et al., 2016) and (Tano et al., 2021) for swimming motility and from (Pelaez-Vico et al., 2016) for swarming motility. Briefly, overnight cultures of *R. solanacearum* strains were grown at 28°C in MP medium supplemented with 10 mM L-glutamate as the sole carbon source (Plener et al., 2010) and used to inoculate new cultures in the same medium with an initial DO_{600} of 0.1. The cultures were grown for 5 hours. Freshly prepared MP medium semi agar (0.3% for swimming motility or 0.6% for swarming motility) supplemented with 10 mM L-glutamate was cooled at 55°C before the test and poured in Petri dishes atthe time of use. After 15 min, 2 µL of bacterial cultures diluted to 5.10^7 CFU/mL were stab-inoculated into the agar plates for swimming motility or inoculated onto the agar surface and left to dry for 4 minutes for swarming motility. The diameters of the swimming and swarming halos were measured after 72 hours of incubation at 28°C. All assays were repeated at least three times independently with 3 technical replicated for each strain per assay. Mutant strains were compared to their respective parental strains using a Wilcoxon test ($P<0.05$) performed in the R statistical environment.

Exopolysaccharide production assays

Exopolysaccharide extraction and quantification were performed using the protocol described by (Peyraud et al., 2017) with some method adaptation steps. Cultures in exponential growth

phase of the *R. solanacearum* strains made in MP medium supplemented with 20 mM L-glutamate were used to inoculate new cultures in MP medium supplemented with 10 mM L-glutamate adjusted to an initial OD_{600} of 0.075 and let grown for 15 hours. EPS were extracted from 800 µl of filtered culture supernatants and amounts of N-acetylgalactosamine were quantified according to the protocol described by (Peyraud et al., 2017). For each bacterial strain, EPS production was measured in three independent experiments. Amounts of N-acetylgalactosamine produced by the mutant strains were divided by the amount produced by the parental strain and ratios were compared to the value 1 using a two-sided *t*-test.

Biolog Phenotyping

Metabolic and stress resistance profiles of each strain were established using Biolog phenotype microarrays. PM01 and PM02 were used for utilization of carbon sources, PM03 was used for utilization of nitrogen sources and PM09 and PM10 were used for osmolyte and pH sensitivity assays using the protocol supplied by the manufacturer (Biolog, Inc., Hayward, CA, USA). All tests were performed during 96 hours at 28°C in an Omnilog reader. Only one set of phenotype microarrays were analyzed per strain. Data were analyzed with the parametric V1.3 Biolog software and the measure of the areas under the curves were used to score the metabolic activities of each strain. Evolved clones carrying the adaptive mutations were compared with their respective ancestors using the package opm in the R environment.

Validation of production of redox energy under acidic conditions

The production of energy by the mutants under acidic conditions observed in Biolog phenotyping microarray assays was validated in plates during 96 hours. To do so, bacteria grown in Petri dishes for 48 hours were resuspended in 10 mL of sterile water, then centrifuged and resuspended in order to have the optical density (DO_{600}) of 0.141 (transmittance of 85%, corresponding to 7.05×10^7 cells/mL) in the solution IF-0 1X provided by the manufacturer. Finally, bacteria were resuspended (1/200 of the final volume of the resuspension) in a mix of IF-10 1X and redox solution 1X, both provided by the manufacturer and a medium MP 1X containing L-glutamate monosodium (10 mM). All plates were incubated in an Omnilog reader at 28°C for 96 hours and analyzed with the parametric V1.3 Biolog software. The measure of the areas under the curves were used to score the metabolic

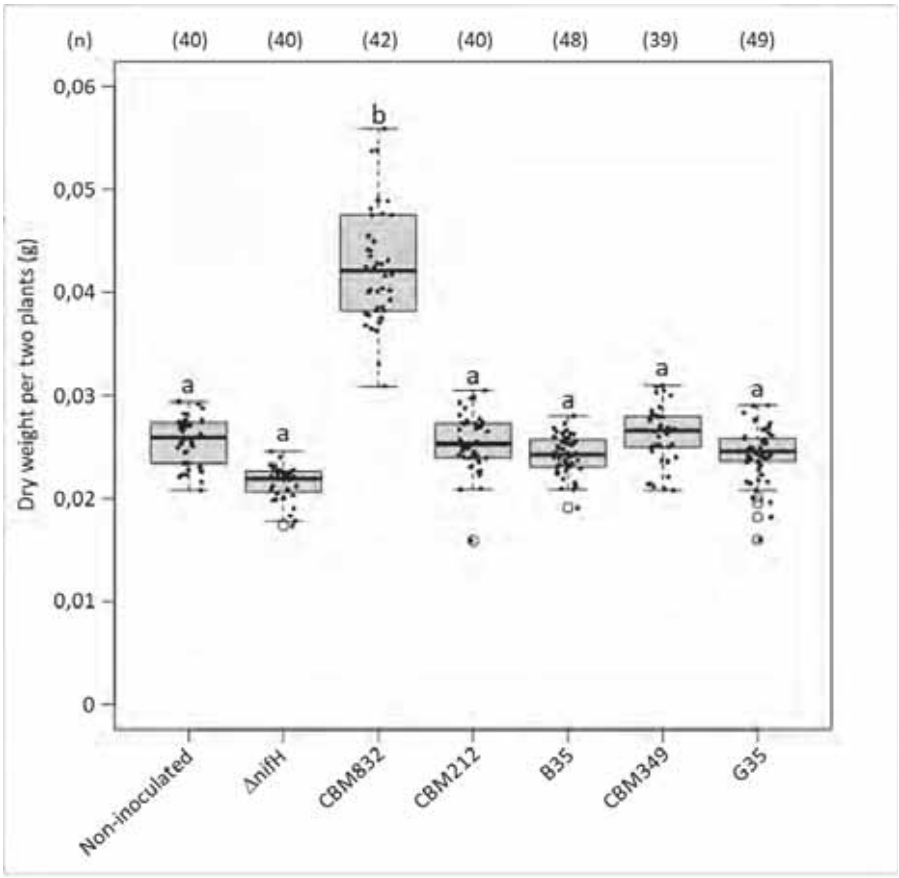


Figure 3.1: Dry weight of *Mimosa pudica* plants inoculated with evolved and natural symbionts. Dry weight of plants 21 days after single inoculation with each strain. Horizontal segments correspond to mean values of dry weight (g). Comparison among groups was made using one-way ANOVA and Tukey post-hoc tests were performed when significant main effects were observed. Means with the same letter do not differ by the Tukey test at 5% probability. Data were obtained from 3 to 4 independent experiments. The sample size (n) is comprised between n=39 – 49. Each value of dry weight was obtained from 2 plants harvested together.

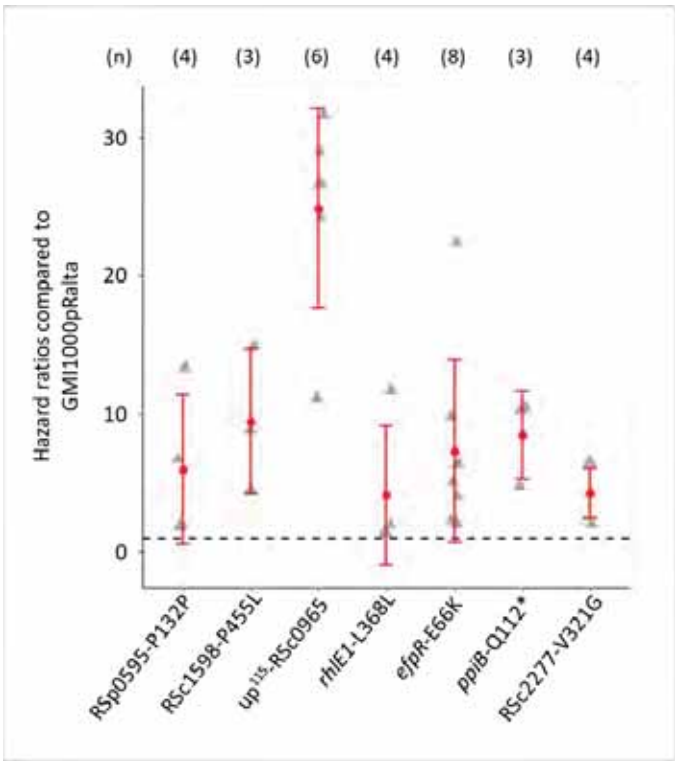


Figure 3.2: Pathogenicity on *Arabidopsis thaliana* of *Ralstonia solanacearum* GMI1000 pRalta mutants carrying highly adaptive mutations for symbiosis. **H**azard ratio was calculated as the time of survival of plants inoculated with the mutants compared with the survival of plants inoculated with the chimeric strain. Red cercles correspond to mean values of Hazard ratios. Red segments indicate the standard deviation of the values from the mean. Gray triangles represent all individual values observed during the assays. Data were obtained from 3 to 8 independent experiments (n). Each value was obtained from pools of 16 plants. Data for *efpR*-E66K and up^{115}-RSc0965 are from Capela et al., 2017.

activities of each strain. Evolved clones carrying the adaptive mutations were compared with their respective ancestors using a one-sided Wilcoxon test (P<0.05) in the R environment.

Results

Symbiotic *Ralstonia* clones did not evolve towards parasitism on *M. pudica*

In the experiment, evolved bacteria were selected for their improved proliferation in *Mimosa* nodules. However, it is not known whether this high level of proliferation incurs a fitness cost for the host plant, in which case they could be considered as parasites. Although *M. pudica* is not a natural host plant of *R. solanacearum,* we wondered whether experimentally evolved *Ralstonia* clones with improved *in planta* fitness on *M. pudica* have evolved towards parasitism on this plant. To answer this question, we inoculated *M. pudica* plantlets individually with the nodulating ancestors and the 35 cycle evolved clones from the two studied lineages (B and G) and measured the plant dry weight 21 days after inoculation. We did not observed any significant differences between the dry weights of plants inoculated with evolved bacteria as compared to non-inoculated plants. This result indicated that nodulation with the non-fixing evolved *Ralstonia* clones is not particularly costly for the plant unlike the interaction with a *C. taiwanensis* non-fixing *nifH* mutant, which significantly reduced the plant dry weights (figure 3.1). We concluded that *Ralstonia* symbionts, although not yet mutualistic, do not behave as parasites on *M. pudica*.

Highly adaptive mutations for symbiosis decrease the pathogenicity of the ancestral *R. solanacearum* strain

Previous results have shown a clear trade-off between symbiotic adaptation with legumes and pathogenicity since all symbio-adaptive mutations characterized so far, including *efpR*-E66K and up^{115}-RSc0965, have decreased or completely abolished the pathogenicity of *R. solanacearum* (Capela et al., 2017; Guan et al., 2013; Marchetti et al., 2010; Tang et al., 2020). To evaluate whether the five newly identified symbio-adaptive mutations affect the pathogenicity of *R. solanacearum*, we reconstructed these mutations individually in the pathogenic ancestral strain *Ralstonia solanacearum* GMI1000pRalta, inoculated them on the model plant *Arabidopsis thaliana,* and followed disease symptoms (wilting) over time on the plants. Four of the five mutants caused delayed symptoms in *Arabidopsis*, although with some variation between experiments (Figure 3.2). Only the *rhlE1*-L368L mutation seems to have no

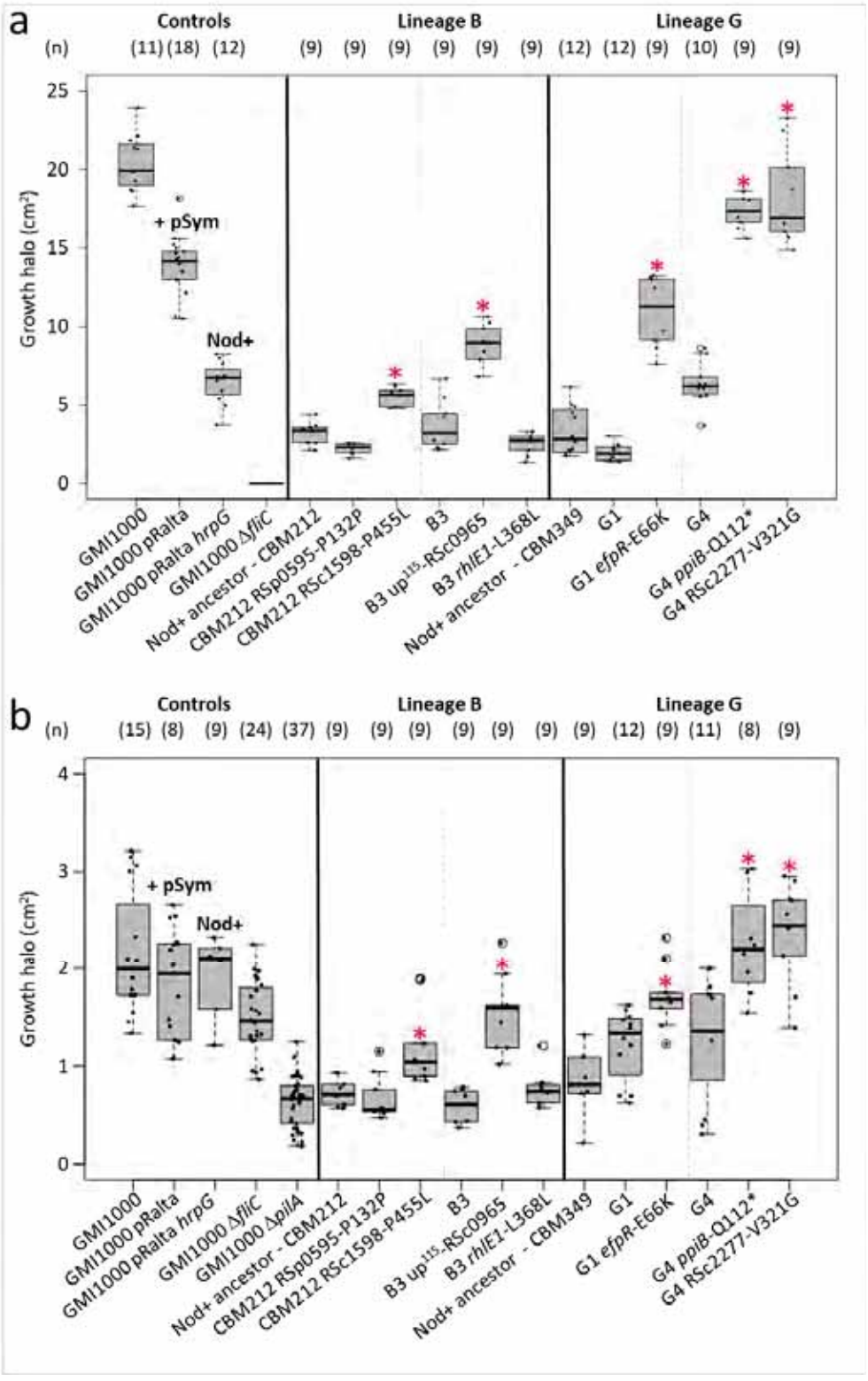


Figure 3.3: Motility of evolved clones carrying highly adaptive mutations for symbiosis. Swimming (a) and swarming (b) motility were measured as the growth halo of bacteria (cm^2) after 48 h of culture in semi-solid media (0.3% and 0.6% of agar, respectively). Horizontal segments correspond to mean values of growth halo. Red stars indicate mutations significantly increasing the bacterial swimming motility (P<0.05, Wilcoxon test). Nod+, nodulating strains. +pSym, chimeric *Ralstonia* strain carrying the symbiotic plasmid from *Cupriavidus taiwanensis*. Data were obtained from 3 to 4 independent experiments. The sample size (n) is comprised between n=9-18. Each value was obtained from one colony.

significant effect on pathogenicity. It is interesting to note that this mutation arose in a genetic background containing the up^{115}-RSc0965 mutation, which confers a quasi-avirulent phenotype in the GMI1000pRalta strain. These results therefore confirmed that adaptation to symbiosis with legumes is globally associated with a loss of bacterial pathogenicity, although there is no quantitative relationship between symbiotic fitness gains and decreases in pathogenicity (Tables 3.1 and 3.2).

Symbiotic adaptation is associated with an increase in bacterial motility and a decrease in EPS production

To characterize the biological functions affected by the symbio-adaptive mutations, we started by measuring their effect on bacterial motility and production of exopolysaccharides (EPS), two traits important for both mutualistic and pathogenic interactions and subject to multiple regulations. In order to take into account the genetic background in which these mutations have arisen and the potential epistatic effects with previously arising mutations, we introduced these mutations individually in their closest ancestral evolved clone (Doin de Moura et al., 2022). Motility and EPS production capacities of evolved clones and their isogenic mutants were compared together with the non-nodulating and nodulating ancestors.

First, swimming and swarming motility strongly decreased in the nodulating ancestors compared to GMI1000. This decrease was mainly due to the acquisition of the symbiotic plasmid pRalta and to the *hrpG* stop mutations allowing nodulation in CBM212 and CBM349 (Figure 3.3). Then, in both lineages, several adaptive mutations (RSc1598-P455L and up^{115}-RSc0965 in lineage B and *efpR*-E66K, *ppiB*-Q112* and RSc2277-V321G in lineage G) increased the swimming and to a lesser extent the swarming capacity of evolved bacteria, and restored either partially (lineage B) or completely (lineage G) the motility level of the ancestral strain GMI1000pRalta (Figure 3.3).

On the other hand, EPS production was drastically reduced in a convergent manner in both lineages. In line with previous findings, mutations affecting the EfpR pathway (up^{115}-RSc0965 and *efpR*-E66K) strongly reduced EPS production while enhancing motility (Capela et al., 2017; Perrier et al., 2016). The *ppiB*-Q112* and Rsc2277-V321G mutations, which have increased motility, also appeared to slightly reduce EPS production (Figure 3.4). However, this reduction was difficult to estimate since these mutations occurred in an *efpR* background, which already produced low levels of EPS. Lower EPS production and better motility were not always

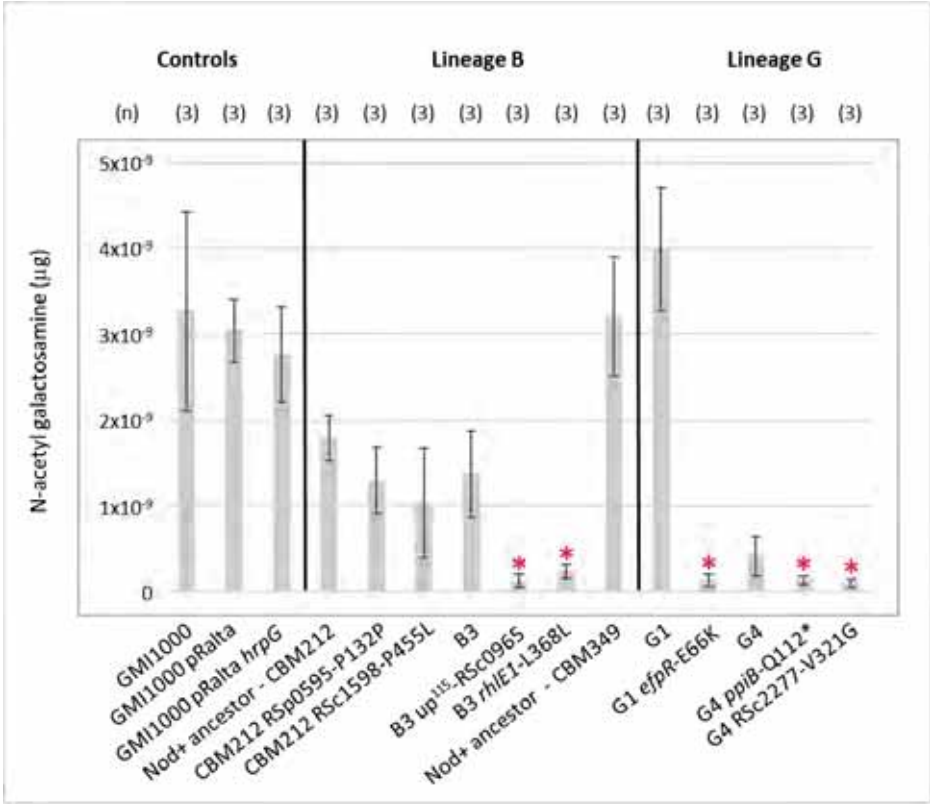


Figure 3.4: Exopolysaccharide production by mutants carrying highly adaptive mutations for symbiosis. EPS production was measured by extracting and quantifying the production of N-acetyl galactosamine by the mutants or parental strains. Bars correspond to mean values +/- standard deviations. Red stars indicate mutations significantly decreasing the EPS production compared to the parental strain (P<0.05, t-test). Data were obtained from 3 independent experiments. Each value was obtained from 1 measure of EPS production via colorimetric assays.

correlated. For example, the nodulating ancestors with lower motility than the non-nodulating ancestor were not affected for EPS production. Similarly, the mutation *rhlE1*-L368L strongly reduced EPS production without affecting bacterial motility.

Metabolic capacities and tolerance to acidic pH were gained in both evolved lineages

Previous results showed that the *efpR*-E66K mutation inactivating the regulator EfpR enhanced the ability of *R. solanacearum* to catabolize a large number of carbon and nitrogen sources (Capela et al., 2017; Perrier et al., 2016). To assess whether the newly found adaptive mutations have any effect on bacterial metabolism, we analyzed the metabolic profiles of evolved clones and their isogenic derivative mutants using Biolog phenotype microarrays dedicated to the use of carbon and nitrogen sources and stress resistance. Expectedly, the clone B3 up^{115}-RSc0965 displayed a similar metabolic profile as the G1 *efpR*-E66K mutant. Interestingly, the mutant B3 *rhlE1*-L368L has an almost identical metabolic profile than the *efpR*-E66K and up^{115}-RSc0965 mutants, with the exception of the Ala-His dipeptide as nitrogen source which is better utilized by the *rhlE1* mutant only (Figure 3.5, Sup Table 3.3). In these three mutants, the catabolism of intermediary components of the TCA cycle, citric acid, fumaric acid, L-malic acid, and succinic acid is enhanced. Several amino-acids are better used either as carbon (L-histidine, L-proline, L-serine, L- and-alanine) or nitrogen (L-tyrosine) sources while the purine base adenine is better used as nitrogen source. Strikingly, five mutations, RSc1598-P455L, RSp0595-P132P, *efpR*-E66K, up^{115}-RSc0965, and *rhlE1*-L368L allowed a higher production of energy in the presence of sodium nitrate, suggesting an activation of the denitrification pathway by these mutations, a process that may be important under oxygen limiting conditions (Fischer, 1994). Altogether, these results showed a large range of compounds whose catabolism was convergently improved in the two lineages (Figure 3.5). Besides, Biolog stress resistance microarrays showed that three mutations, up^{115}-RSc0965 and *rhlE1*-L368L in lineage B and RSc2277-V321G in lineage G, conferred the ability to produce energy at pH5. This better tolerance to acidic pH was statistically confirmed only for the mutation RSc2277-V321G by custom-made Biolog plates (Figure 3.6). We also tested the better tolerance to acidic conditions of the other adaptive mutants as well as the final cycle 35 evolved clones using the same custom-made Biolog plate and found that the mutants CBM212 RSc1598 – P455L and G1 *efpR*-E66K also showed an improved production of energy compared to their respective parental strain at pH5 (P<0.05, one-sided Wilcoxon test). We

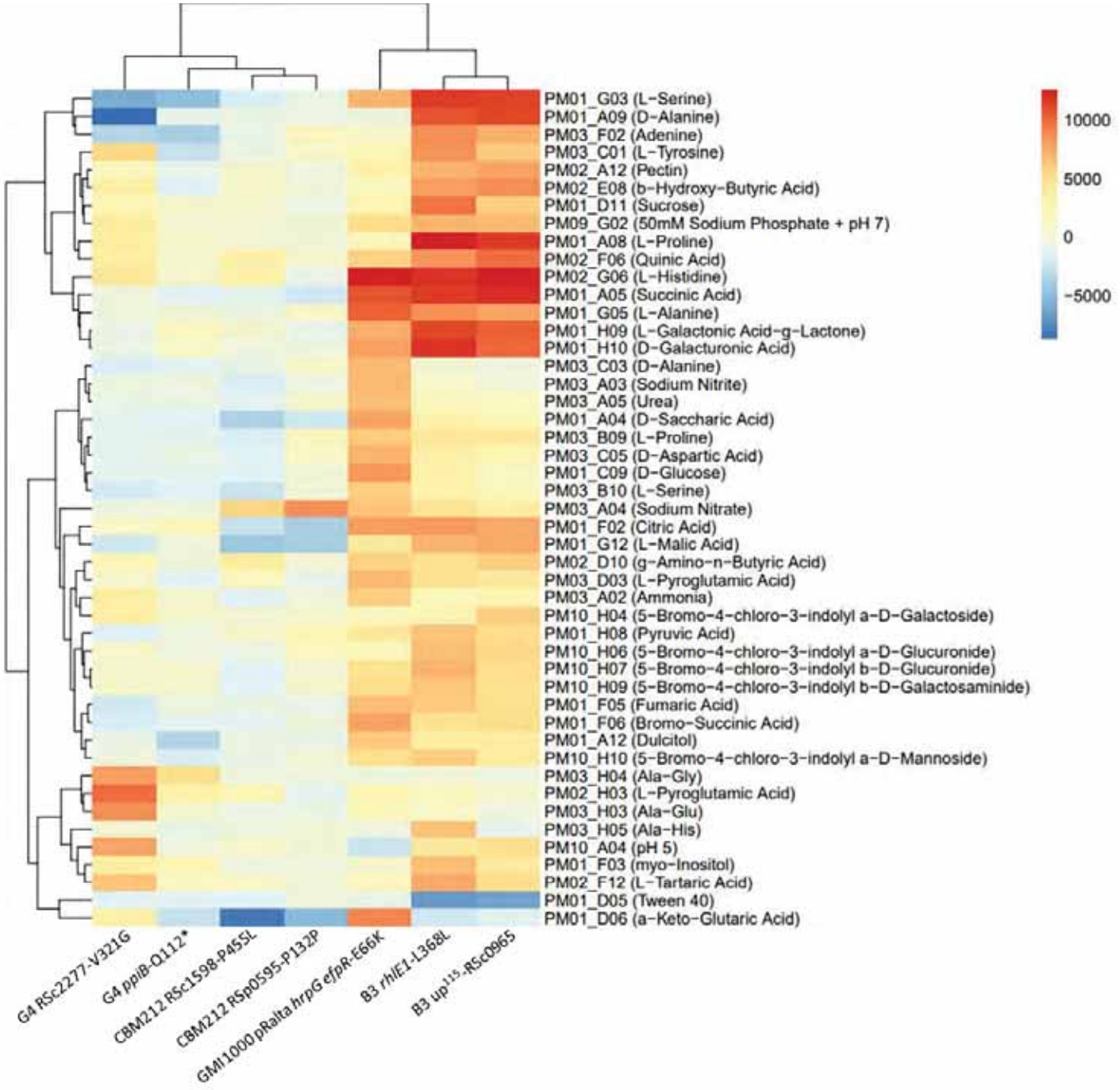


Figure 3.5 – Effect of the seven highly adaptive mutations for symbiosis on the metabolic profile of the evolved clones. Measurements of production of redox energy *via* Biolog phenotyping microarray assays. Values represent the difference between the area under the curve of the mutant and the parental strain (cut-off threshold of 6000). Redox energy production changes are represented in the heatmap by variations between blue (weakest values) and red (highest values). PM01, carbon source. PM02, carbon source. PM03, nitrogen source. PM10, stress tolerance.

also tested whether the tolerance to acidic pH could be related to a reduction of EPS production by testing an *epsA* mutant using a custom-made Biolog plate. However, this mutant did not allow bacteria to produce energy at pH5 indicating that the reduction of EPS production is not responsible for acid stress tolerance in evolved bacteria (Figure 3.6).

Discussion and perspectives

In this study, we characterized seven highly adaptive mutations for symbiosis with the objective to identify the biological functions involved in the transformation of the plant pathogen *R. solanacearum* into a legume symbiont. Previous work showed that these mutations strongly improved the first stages of symbiosis, particularly the ability of bacteria to compete for nodule occupancy and, for some of them, the ability to proliferate within nodules (Doin de Moura et al., 2022). However, once inside the host cells, all evolved clones degenerate prematurely, preventing the establishment of a mutualistic interaction with the plant (Doin de Moura et al., 2022 and unpublished data). It is important to note that at that stage of the study, we do not know whether the symbiotic adaptation relied on a loss or a gain of function of the mutated genes, except for *efpR*-E66K and up^{115}-RSc0965 that both repress *efpR* expression (Capela et al., 2017) and the stop mutation in *ppiB* that likely alters the function of the corresponding protein. This will have to be determined by deleting the mutated genes and testing their symbiotic fitness.

First, to characterize these mutations, we evaluated their effect on the virulence of the ancestral pathogenic strain GMI1000 pRalta. Results highlighted a strong antagonism between pathogenicity and symbiotic adaptation, since all but one of the mutations delayed disease symptoms on *A. thaliana.* This suggests that weakly immunogenic bacteria are probably preferentially selected by the plant and allowed to enter the root. Indeed, expression of plant defense genes were shown to be activated in many legumes few hours after inoculation with both compatible and incompatible rhizobia, but then rapidly suppressed only in the presence of compatible symbionts (Gourion et al., 2015; Jones et al., 2008; Kouchi et al., 2004; Libault et al., 2010). A stronger –or longer- activation of defense responses in ancestral bacteria may thus be detrimental to symbiosis.

By exploring a number of biological functions, we found that most of the symbio-adaptive mutations have pleiotropic effects. Indeed, six of the seven mutated genes are putatively involved in (global) regulation either at the transcriptional (RSp0595, RSc1598, *efpR*, RSc0965),

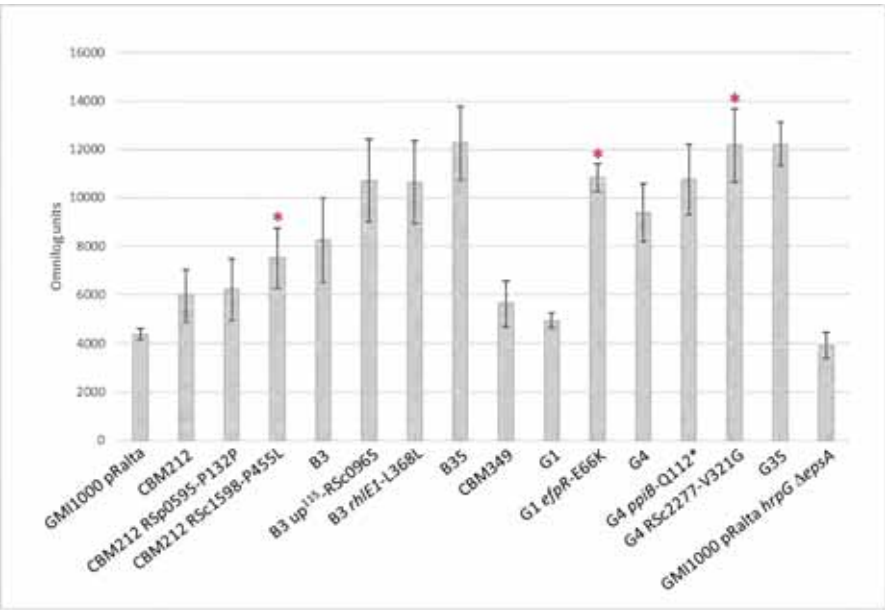


Figure 3.6 – Production of redox energy by the mutants under acidic conditions. Measurement of the redox energy produced by the mutants and their respective parental strains in an acidic medium (MP L-glutamate monosodium 20mM, pH 5.0). Red stars indicate significant increases in the energy production of the mutant compared to its parental strain (P<0.05, one-sided Wilcoxon test).

post-transcriptional (*rhlE1*) (Hausmann et al., 2021) or translational (*ppiB*) (Skagia et al., 2016) levels. Among the functions modified by the mutations, swimming motility, EPS production, metabolism, and some stress tolerance were convergently affected in the two lineages studied, suggesting that these biological traits could be involved in symbiotic adaptation.

Swimming motility strongly fluctuated during the course of the evolution experiment, being first markedly reduced by the transfer of the symbiotic plasmid *into R. solanacearum* and the acquisition of the adaptive mutations in *hrpG* conferring nodulation capacity, then being progressively increased over evolution cycles. However, evolved clones and reconstructed mutants have not shown higher swimming capacities than the ancestral strain GMI1000. Improvements in bacterial swimming motility could be particularly important in early steps of the symbiosis, allowing a better access to the host root and/or to nutrient sources in the rhizosphere. Experimental evolution of *Pseudomonas protegens* showed that motility is one of the crucial function for bacteria to colonize a new host rhizosphere (Li et al., 2021). Motility was also shown to be determinant for nodule initiation and occupancy in several rhizobium-legume interactions (Caetano-Anollés et al., 1988; Zdor & Pueppke, 1990). Moreover, transposon-insertion sequencing provided evidences that bacterial motility also plays an important role during nodule infection and bacteroid formation (Wheatley et al., 2020). However, the importance of motility at later stages of the symbiosis is still an open question since transcriptomics have showed a global repression of flagellin, chemotaxis and motility genes during symbiosis (Tambalo et al., 2015; Uchiumi et al., 2004). The absence of flagella in bacteroids has also been demonstrated (Tatsukami et al., 2013; Uchiumi et al., 2004) but other transcriptomic analyses associated with laser microdissection of nodule tissues showed a reactivation of flagellin genes in the nitrogen fixing zone of the nodules (Roux et al., 2014).

Besides bacterial motility, exopolysaccharides (EPS) also play an important role for both symbiotic and pathogenic interactions. Among the seven adaptive mutations tested, five decreased this trait. Rhizobial mutants impaired in EPS biosynthesis have been shown to be ineffective symbionts forming abortive infection threads (Cheng & Walker, 1998a; Maillet et al., 2020). Indeed, these molecules are specifically recognized by a plant receptor able to distinguish between compatible and incompatible bacterial EPS structures during the infection process (Kawaharada et al., 2015). On the other side, EPS are also important for the pathogenicity of *R. solanacearum*. In this bacterium, EPS are used to hide superficial features that could be recognized by the host plant and activate plant immunity. In addition,

Table 3.2. Functional characterization of the mutants

Mutation[a]	Hazard ratio[b]	Improvement on swimming motility[c]	Effect on EPS production[d]	Tolerance to acidic pH[e]	Metabolic gains		
					Number of new carbon sources	Number of new nitrogen sources	Total number of new sources
RSp0595-P132P	5.98	NE	NE	NE	0	1	1
RSc1598-P455L	9.41	+	NE	NE	0	1	1
up^{115}-RSc0965	24.87	+	-	+	24	17	41
rhlE1 -L368L	4.12	NE	-	+	24	18	42
efpR -E66K	7.31	+	-	+	21	20	41
ppiB -Q112*	8.49	+	-	+	4	1	5
RSc2277-V321G	4.27	+	-	+	9	11	20

[a]up^{x}, intergenic mutations located x nucleotides upstream the start codon of the gene indicated. [b]Mean values of hazard ratio compared to the chimeric strain GMI1000 pRalta obtained from 3 to 8 independent experiments. [c]+, improvement of bacterial swimming motility caused by the adaptive mutation. [d]-, decrease in the amount of produced exopolysaccharides caused by the adaptive mutation. [e]+, increase in the tolerance of the evolved clone to acidic pH. [c,d,e] NE, no effect identified of the mutation on the evaluated trait.

overproduction of EPS molecules in the xylem block the water traffic in these vessels and trigger plant wilting. In this scenario, we thus have an apparent paradox between the requirement of EPS for both symbiosis and pathogenicity and the reduction of EPS synthesis by the symbio-adaptive mutations. However, since *R. solanacearum* and *C. taiwanensis,* the natural *Mimosa* symbiont, possess completely different sets of EPS biosynthesis genes (Amadou et al., 2008), it is possible that EPS produced by *Ralstonia* are not (or not completely) compatible with the *Mimosa* receptor and impair the infection process. Besides, in previous work, we showed that inactivation of the EPS regulator XpsR in the nodulating strain GMI1000 pRalta *hrpG* did not improve intracellular infection of *M. pudica* nodules (Tang et al., 2020), suggesting that either reduction of EPS production is not involved in symbiotic adaptation or that it is not sufficient to explain alone the adaptation.

The rhizobium-legume nitrogen-fixing symbiosis is a nutritional symbiosis, which involves multiple exchanges of carbon and nitrogen molecules between the bacterial and plant partners. The ability of bacteria to assimilate available nutrient sources both in the rhizosphere and inside the host plant is crucial for bacterial survival and for the success of these mutualistic interactions. Therefore, we looked for metabolic adaptations conferred by the symbio-adaptive mutations. First, we observed that all the mutations conferred metabolic gains and no losses. Strikingly, three mutations, *efpR*-E66K, up^{115}-RSc0965 and *rhlE1*-L368L, deeply modified the metabolic capacities of bacteria. Noteworthy, these three mutations improved the rhizosphere colonization, nodule occupancy and within-nodule proliferation capacities of bacteria. Moreover, these three mutants displayed a very similar metabolic profile. This was expected for *efpR*-E66K and up^{115}-RSc0965 mutants (Capela et al., 2017) but not for the *rhlE1*-L368L mutant, suggesting a possible connection between *rhlE1* and the *efpR* pathway. These three mutations allowed a better use of a large range of carbon and nitrogen compounds (Figure 3.5), indicating that these genes are involved in global metabolic regulation. Alternatively, the enhanced metabolism in these mutants could be due to a reallocation of resources following the decrease in the production of costly virulence factors such as EPS (Peyraud et al., 2016). Among their metabolic changes, some intermediates of the tricarboxylic acid (TCA) cycle (pyruvic acid, citric acid, succinic acid, fumaric acid and malic acid) were better used. This is particularly interesting, since some of those TCA intermediates are also products of photosynthesis, being possibly provided by the plant host. Indeed, the role of those molecules in the nitrogen-fixing symbiosis have been well described (van Slooten

et al., 1992; Yurgel and Kahn, 2005). For example, C4-dicarboxylate transport mutants of *Rhizobium trifolii* unable of assimilating and transporting succinate, malate and fumarate but able to use several other sources of carbon formed ineffective nodules in both white and red clover (C. Ronson et al., 1981). Nevertheless, evolved clones appearing in this evolution experiment do not form functional nodules, leading us to think that improvement of C_4-dicarboxylate metabolism may not be responsible for the adaptation of evolved clones to the first stages of symbiosis.

Additionally, several amino acids were better used by the mutants either as carbon (L-proline, L-serine, L-alanine, L-histidine) or nitrogen (L-tyrosine, D-aspartic acid, L-serine, D-alanine and L-pyroglutamic acid) sources in both lineages. Better use of amino acids was shown to be an essential trait enhancing the competitiveness for nodule formation of the soil-borne bacterial strains (Wielbo et al., 2007). Reinforcing these findings, genes related to the catabolism of amino acids were shown to be upregulated in bacteroids (R.T. Green et al., 2019), while biosynthesis genes of some amino acids, such as the branched chain amino acids, were shown to be downregulated (Prell et al., 2009). Moreover, auxotrophic mutants unable to biosynthesize amino acids have been shown to present a Fix+ phenotype on some legume plants, suggesting that in these cases plants provide amino acids to the bacteria during the symbiosis (Dunn, 2014; J.I. Jiménez-Zurdo et al., 1995; Kumar et al., 2003). The proline amino acid seems to be particularly important during the nitrogen-fixing interaction (diCenzo et al., 2015). Indeed, transcriptomics approaches have suggested that the ability to assimilate proline as source of carbon by strains of *Bradyrhizobium* were crucial during the early steps of symbiotic interaction with soybean after the exposition of the bacterial strain to soybean exudates (Liu et al., 2017), and rhizobium deletion mutants unable to catabolize proline were shown to be defective for induction of nodule organogenesis in soybean and for nitrogen fixation (Sabbioni and Forlani, 2022). Interestingly, some D-amino acids were better used by the symbio-adaptive mutants. Bacteria predominantly use L-amino acids for basal metabolism and reconstruction of cell component but D-amino acids were shown to have regulatory role, in particular in the cell wall remodeling during bacterial stationary growth phase. These amino acids may be important for a better resistance of the peptidoglycan molecules, consequently contributing to the preservation of the cell structure and resistance to environmental conditions (Lam et al., 2009).

Additionally, the mutants *efpR*-E66K, up[115]-RSc0965, and *rhlE1* enabled bacteria to better use purines as nitrogen sources. Nodules of auxotroph mutants of rhizobia unable to assimilate purines have been shown to be small and ineffective (Newman et al., 1992; Stevens et al., 2000). In the same work, the addition of 5-imidazole-4 carbaxamide riboside, the precursor of inosine, restored the nodulation of the auxotroph strains. The observed better assimilation of purines may be particularly advantageous during the symbiotic interaction since these molecules seem to be in higher amount in the nodules compared to other plant organs. An elevated activity (35 fold) of the purine nucleosidase enzyme was specifically found inside nodules of soybean, compared to the concentration of the same enzyme in other plant organs, such as plant roots (Larsen and Jochimsen, 1987). In the same way, purine oxidative enzymes, such as xanthine dehydrogenase and uricase, were isolated in large amounts inside nodules of soybean, indicating that purines are actively catabolized by bacteria in nodules (Reynolds et al., 1982).

Interestingly, five adaptive mutations enabled the bacteria to better use sodium nitrate as source of nitrogen. One adaptive mutation, *efpR*-E66K, also improved the use of sodium nitrite as source of nitrogen. Nitrate and nitrite are important for different metabolic processes in rhizobia, such as the assimilation of nitrate, a highly energy demanding pathway (Dalsing and Allen, 2014), and the denitrification pathway, in which nitrate is used as electron acceptor to produce energy (Dalsing et al., 2015). As the Biolog phenotyping method relies on the production of redox energy, we could hypothesize that a better use of nitrate by those mutants reflects the use of nitrate as final acceptor of electrons during the process of anaerobic respiration *via* denitrification. This could be particularly important during the experimental evolution of *R. solanacearum*, since low levels of oxygen in our experimental system are present both outside the plant, due to the cultivation of *Mimosa* plants in liquid medium, and inside the nodules. Additionally, a better use of sodium nitrate was also the only metabolic change observed for the mutations in RSp0595 and RSc1598. RSp0595 encodes a putative quorum quenching enzyme (a putative AHL-lactonase), which may regulate the level of AHLs quorum sensing molecules, which in turn might activate NarL, a response regulator of the LuxR family located next to the *narGHJI* denitrification genes. RSc1598 encodes a putative oxygen sensing histidine kinase whose target genes in *R. solanacearum* are completely unknown.

Besides that, bacterial stress resistance profiling also evidenced that a better tolerance to acidic conditions was conferred by three mutations (CBM212 RSc1598 – P455L, G1 *efpR*-E66K and G4 RSc2277-V321G). A better tolerance to acidic conditions (pH around 3.5-4.5) was suggested to be a predisposal trait for symbiotic association with legumes (Fagorzi et al., 2020). Indeed it could represent an advantage in both legume rhizosphere and nodule cells. Due to the presence of organic acids in the root exudates, as well as protons related to the maintenance of the net charge across the root membrane (Jones et al., 2003; Moreira and Siqueira, 2006), the rhizosphere has a pH decreased by 2 units compared to the neighboring soil (Faget et al., 2013). Acidic conditions probably also prevail in the nodules due to the activity of H(+)-ATPases in the symbiosome membrane, leading to an accumulation of protons in the peribacteroid space (Hinde and Trautman, 2001; Whitehead and Day, 1997). Consistently, several enzymes working in the symbiosome have an acidic optimal pH (Mellor, 1989; Panter et al., 2000). Bacteria have developed diverse mechanisms to improve their acid tolerance during adaptation to acid environments (Guan and Liu, 2020). Among them, metabolic regulations may play an important role. For example, the neutralization of protons by the production of ammonia and carbon dioxide and the consumption of intracellular H^+ via deamination and decarboxylation of amino acids have been identified as key mechanisms in bacterial resistance to acidification (Lu et al., 2013; Xiong et al., 2014). Accumulation of polyamines may also allow a better survival in acidic pH (Fujihara and Yoneyama, 1993). Ornithine and citrulline, better used by the *efpR*-E66K, up^{115}-RSc0965 and *rhlE1*-L368L mutants, are intermediate compounds of this pathway. Acid stress tolerance was recently reported in *R. solanacearum* mutated in *phcA*, a gene encoding a global regulator of virulence and metabolism (Liu et al., 2022). Interestingly, the *phcA* regulon and metabolic profile largely overlap with those of the *efpR* mutants (Perrier et al., 2018; Peyraud et al., 2016).

Functional convergences observed in this work have provided some clues to identify the biological functions that have possibly contributed to the transformation of *R. solanacearum* in an intracellular legume symbiont. It will now be necessary to evaluate the effect of each of these biological functions individually on symbiotic adaptation. The inactivation of key genes essential for each biological functions coupled to symbiotic phenotyping (relative *in planta* fitness) should allow addressing this question. However, our assays gave us only a partial view of the functions affected by the mutations. Transcriptomic (and proteomic) analyses may be required to get a more complete and exhaustive view of functional changes. The diversity of

functions modified by the symbio-adaptive mutations led us to wonder whether it is necessary to change multiple biological functions at the same time to allow symbiotic adaptation. This hypothesis is relevant, since the nitrogen-fixing interaction implies multiple steps. In each step, the plant host and the environment exposed the bacterial partner to various conditions, very often occurring at the same time. For example, traits related to a better motility, tolerance to acidic pH and a better use of new sources of carbon and nitrogen could be especially important in the rhizosphere. In the same way, tolerance to acidic pH, use of new carbon and nitrogen sources and better use of nitrate could be especially important in the nodule environment, where very specific carbon sources are available and, at the same time, acidic conditions and low levels of oxygen prevails.

Sup Table 3.1. Strains used in this study

Resource type	Designation	Reference/source	Identifier	Additional information/Common name
C. taiwanensis	Wild-type strain isolated from *Mimosa pudica* in Taiwan	Chen et al., 2001	LMG19424	
C. taiwanensis	LMG19424 derivative resistant to StrR	M. Hynes	CBM832	
R. solanacearum	Non-nodulating chimeric ancestor, TriR	Marchetti et al., 2010	CBM124	GMI1000 pRalta::TriR
R. solanacearum	Chimeric ancestor carrying the mutation *hrpG* -Q81* from the nodulating ancestor CBM212, TriR	Marchetti et al., 2010	CBM1627	CBM124 *hrpG* -Q81*
R. solanacearum	Nodulating ancestor mutated in *hrpG* , TriR, GenR	Marchetti et al., 2010	CBM212	CBM212
R. solanacearum	Nodulating ancestor mutated in *hrpG* , TriR, GenR	Marchetti et al., 2010	CBM349	CBM349
R. solanacearum	Evolved clone from line B, cycle 35, TriR, GenR	Doin de Moura et al., 2022	RCM2458	Clone B35
R. solanacearum	Evolved clone from line G, cycle 35, TriR, GenR	Doin de Moura et al., 2022	RCM2460	Clone G35
R. solanacearum	Evolved clone from line B, cycle 3, TriR, GenR	Marchetti et al., 2010	CBM510	Clone B3
R. solanacearum	Evolved clone from line G, cycle 1, TriR, GenR	Marchetti et al., 2010	CBM420	Clone G1
R. solanacearum	Evolved clone from line G, cycle 4, TriR, GenR	Marchetti et al., 2010	CBM557	Clone G4
R. solanacearum	Nodulating ancestor CBM212 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	Marchetti et al., 2017	RCM1833	CBM212 IGglmS:P*psbA* -GFP
R. solanacearum	Evolved clone B3 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3343	Clone B3 IGglmS:P*psbA* -GFP
R. solanacearum	Evolved clone G1 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM1818	Clone G1 IGglmS:P*psbA* -GFP
R. solanacearum	Evolved clone G4 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM1958	Clone G4 IGglmS:P*psbA* -GFP
R. solanacearum	CBM212 Rsp0595-P132P carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3130	CBM212 Rsp0595-P132P IGglmS:P*psbA* -mCherry
R. solanacearum	CBM212 RSc1598-P455L carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM2166	CBM212 RSc1598-P455L IGglmS:P*psbA* -GFP
R. solanacearum	Evolved clone B3 up^{115}-RSc0965 carrying a P*psbA* -GFP fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3157	B3 up^{115}-RSc0965 IGglmS:P*psbA* -GFP
R. solanacearum	Evolved clone B3 *rhlE1* -L368L carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3276	B3 *rhlE1* -L368L-IGglmS:P*psbA* -mCherry
R. solanacearum	Evolved clone G1 *efpR* -E66K carrying a KanR cassette downstream glmS, TriR, KanR	Capela et al., 2017	RCM1867	G1 *efpR* -E66K IGglmS:KanR
R. solanacearum	Evolved clone G4 *ppiB* -Q112* carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3084	G4 *ppiB* -Q112* IGglmS:P*psbA* -mcherry
R. solanacearum	Evolved clone G4 RSc2277-V321G carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3083	G4 RSc2277-V321G IGglmS:P*psbA* -mcherry
Plasmid	Symbiotic plasmid of LMG19424 (0.5 Mb), likely auto-transferable	Amadou, et al. 2008	pRalta	
Plasmid	Plasmid for *R. solanacearum* chromosomal integration of the constitutive *psbA* promoter fused to GFPuv into the intergenic region downstream *glmS* , KanR	Capela, et al. 2017	pRCK-P*ps* -GFP	
Plasmid	Plasmid for *R. solanacearum* chromosomal integration of the constitutive *psbA* promoter fused to mCherry into the intergenic region downstream *glmS* , KanR	Capela, et al. 2017	pRCK-P*ps* -mCherry	

TriR, trimethoprim resistance. KanR, kanamycin resistance. StrR, streptomycin resistance. IGglmS: insertion in the intergenic region downstream the *glmS* gene.

CHAPTER 4

Continuation of the evolution experiment did not lead to mutualism but evidenced a strong mutational convergence in the RSc2277 gene

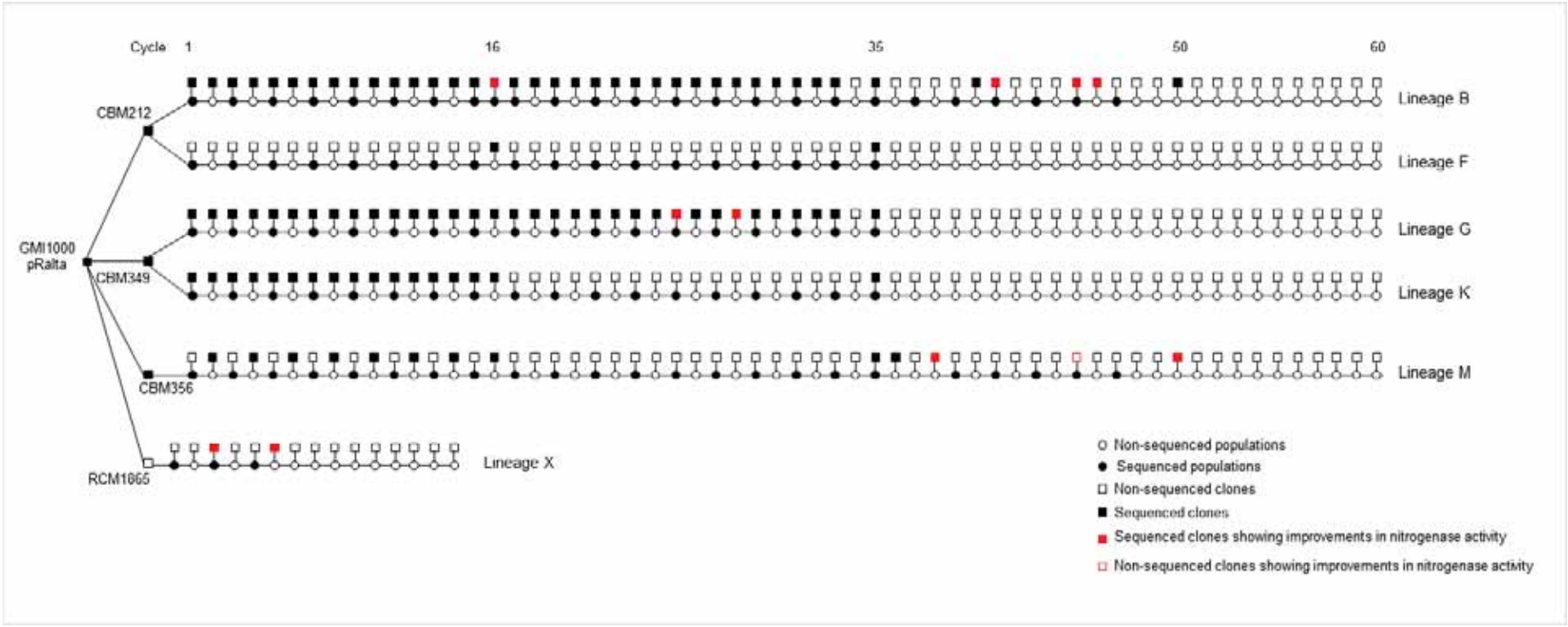


Figure 4.1. **Experimental evolution of *Ralstonia solanacearum* into *Mimosa pudica* symbionts during 60 cycles**. Bacteria were evolved in cycles of inoculation on *M. pudica* plants, harvest of nodule bacteria 21 days after inoculation and re-inoculation of the evolved populations on new plants. Five lineages were evolved in addition to one lineage, which was experimentally evolved from the reconstructed strain GMI1000 pRalta White circles indicate non-sequenced evolved populations; black circles indicate sequenced populations. White squares indicate non-sequenced clones and black squares indicate sequenced evolved clones. Red squares indicate sequenced clones showing increased levels of nitrogenase activity. Non-filled red squares indicate non-sequenced clones showing increased levels of nitrogenase activity.

Abstract

During the experimental evolution of the phytopathogen *Ralstonia solanacearum* into *Mimosa pudica* symbionts, 35 evolution cycles allowed the rapid improvement of early symbiotic traits, in particular nodulation competitiveness (*i.e*. ability to form nodules) and within-host proliferation (*i.e*. ability to infect and multiply in nodules) of *Ralstonia* evolved clones. However, mutualism was not obtained. In order to further evolve bacteria and complete the symbiotic process, additional evolution cycles were performed by the team. After 60 nodulation cycles, representing ca. 1500 bacterial generations, mutualism has still not been observed but low levels of nitrogenase activity were detected in several lineages. The sequencing of evolved populations between the cycles 35 and 60 allowed the identification of a gene, RSc2277, convergently mutated in most of them. Mutations in this gene increased the infectivity of bacteria and their nitrogenase activity although the latter phenotype was only observed in an *efpR* mutated background. After 60 evolution cycles, infectivity and intracellular persistence did not reach the level of the natural symbiont, *Cupriavidus taiwanensis*, in this experiment, potentially because some important late symbiotic functions are absent in *R. solanacearum,* thus preventing the realization of mutualism.

Introduction

The experimental evolution of the phytopathogen *Ralstonia solanacearum* into a symbiont of the legume *Mimosa pudica* rapidly generated bacteria capable of carrying out the first steps of the symbiotic interaction (Doin de Moura et al., 2022; Marchetti et al., 2010; Marchetti et al., 2014). During the first 35 cycles, nodulation competitiveness was optimized in several lineages and has reached the level of the natural *Mimosa* symbiont *C. taiwanensis*. By contrast, the level of infectivity (the number of viable bacteria per nodule), although greatly increased during evolution cycles, did not reach that of *C. taiwanensis* (Doin de Moura et al., 2022). Moreover, mutualism was not achieved since plants did not benefit from the association with evolved bacteria for their growth (Doin de Moura et al., 2022). These observations prompted us to hypothesize that additional evolution cycles may be necessary to complete the symbiotic process. Indeed mutualistic nitrogen fixation is a complex symbiotic stage, which not only relies on the capacity of bacteria to reduce atmospheric nitrogen into

ammonia and deliver it to the plant but also on their capacity to survive and persist within host cells. Persistence of bacteria in nodule cells is a prerequisite for nitrogen fixation that, we thought, could be improved through more nodulation cycles.

In this study, we further evolved the *Ralstonia* populations through 25 additional cycles until the cycle 60 and we started a new lineage from a *Ralstonia* strain in which we introduced two main adaptive mutations allowing nodulation (Marchetti et al., 2010) and good intracellular infection (Capela et al., 2017). We showed that these additional cycles did not lead to mutualism. However, low levels of nitrogenase activity were detected in four evolved lineages. The sequencing of evolved populations evidenced that shifts in nitrogenase activity were associated with mutations targeting one gene, RSc2277, encoding a putative transporter, convergently mutated in all nitrogen-fixing evolved lineages.

Results

Continuation of the evolution experiment led to the detection of low levels of nitrogenase activity in four lineages

Since mutualism was not achieved in the B, F, G, K and M lineages after 35 cycles (Doin de Moura et al., 2022), the team pursued the evolution of these lineages for 25 additional evolution cycles until cycle 60. To reduce the heterogeneity of symbiotic phenotypes due to the heterogeneity of plant genotypes, they used *M. pudica* seeds that were genetically more homogeneous, produced from a single plant clonally propagated by cuttings. Another potential limitation to obtain mutualistic nitrogen fixation was the accumulation of many mutations all along the experiment, some of which being potentially deleterious for late stages of symbiosis and nitrogen fixation. For example, many mutations on the *nif* and *fix* genes, that are essential for nitrogen fixation, occurred in this experiment. Therefore, the team evolved in parallel a new lineage, the lineage X, starting with the GMI1000 pRalta strain carrying the mutation allowing nodulation (*hrpG**) and the highly adaptive mutation *efpR*-E66K improving intracellular infection (Capela et al., 2017). This strain, which was already well adapted for symbiosis, allowed to start evolution in a clean genetic background free of potentially deleterious mutations for symbiosis. This new lineage was evolved for 15 nodulation cycles (Figure 4.1).

At the end of the evolution experiment, none of the plants appeared to be greener, showing that plants did not benefit from nitrogen fixation. To assess whether nitrogenase was active in evolved clones, we performed acetylene reduction assays (ARA) on nodules induced by some evolved clones from the 6 lineages.

No nitrogenase activity was detected in the F and K lineages. By contrast, detectable nitrogenase activities were measured in the four other lineages (B, G, M and X) (figures 4.2a, 4.2c, 4.2e and 4.2f, respectively), although these activities were very low, representing less than 1% of the activity measured in nodules induced by *C. taiwanensis* (figure 4.2g), and transient, the maximum being reached at 15 days post-inoculation (dpi) and significantly decreased at 21 dpi (data not shown). Moreover, these activities were not stably maintained along the lineages but varied strongly between clones from consecutive cycles. In the lineage B, a first level of nitrogen fixation was detected before cycle 16. This level improved after cycle 40 to reach a maximum in clones B41, B45 and B46 then dropped to 0 in clone B50 and was very low in clone B60 (figure 4.2a). In the lineage G, a first level of nitrogenase activity was detected before cycle 25, reached a maximum at cycle 28, was maintained to a lower level until cycle 55 and then almost lost at cycle 60 (figure 4.2c). In the lineage M, a first level of nitrogenase activity was observed at the cycle 38, then this level was more or less maintained until the cycle 60 (figure 4.2e). Finally, clones from the lineage X rapidly acquired the capacity to fix low levels of nitrogen. The maximum was reached in clone X6 and then decrease until cycle 15 (figure 4.2f).

Mutational convergence in the RSc2277 gene is associated with low levels of nitrogen fixation

To analyze the genetic bases of nitrogen fixation shifts, we used previously generated sequencing data of evolved populations and clones for shifts occurring before cycle 35 (Doin de Moura et al., 2022) or we sequenced newly evolved populations and clones. For the lineage G, we previously identified 13 adaptive mutations improving the symbiotic capacity of bacteria between cycle 1 and cycle 35. We assessed the involvement of 5 of them, which were linked to shifts in within-host proliferation, in nitrogen fixation improvements using ARA on nodules formed by the reconstructed mutants described previously (Doin de Moura et al., 2022) and their respective parental strains (Figure 4.3). We observed that the mutations RSc2277-V321G

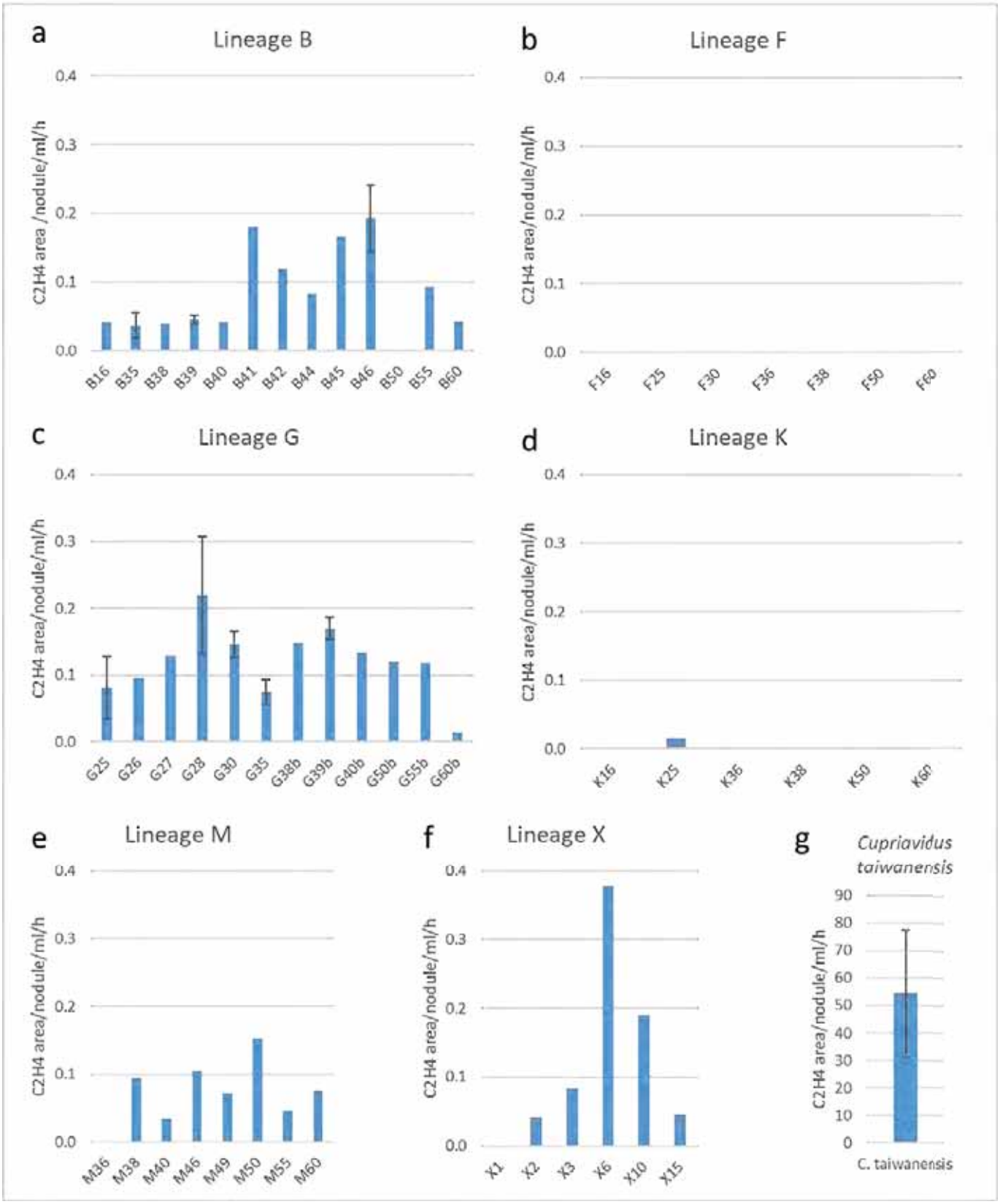


Figure 4.2. **Nitrogenase activity of the evolved clones from the lineages B (a), F (b), G (c), K (d), M (e), X (f) and the natural symbiont of *Mimosa pudica*, *Cupriavidus taiwanensis* (g)**. Data represent the reduction of acetylene per nodule at 15 days post inoculation. Six plants were pooled, exposed to an excess of acetylene during 4 hours. Ethylene amounts were measured through gas chromatography. Data were obtained from a single experiment or up to 8 independent experiments (G25 and G28).

Table 4.1. Candidate cohorts linked to shifts in nitrogenase activity

Cohort	Position	Mutation[a]	Annotation	Gene[b]	Description	Clone
CoB17	1440423	2 bp-AA	Coding (3-4/1746 nt)	*ilvG*	Acetolactate synthase 2 catalytic subunit	B41
CoB17	966048	C-G	I20M (ATC-ATG)	*rfaE*	ADP-heptose synthase protein	B41
CoB17	1708690	G-A	F118F (TTC-TTT)	*xthA*	Exodeoxyribonuclease III protein	B41
CoB17	340629	2 bp-TT	Coding (2328-2329/2811 nt)	pRALTA_0394	Fused transposase IS66/IS21	B41
CoB17	140835	C-T	T93T (ACC-ACT)	RSc0118	Hypothetical protein	B41
CoB17	913640	A-T	I113F (ATC-TTT)	RSc0876	Hypothetical protein	B41
CoB17	1731586	G-A	L819L (CTG-CTA)	RSp1374	Hypothetical protein	B41
CoB17	1991402	C-T	R17W (CGG-TGG)	RSp1586	Hypothetical protein	B41
CoB17	2473740	T-A	L205Q (CTG-CAG)	RSc2277	Hypothetical protein	B41
CoB17	3520496	A-T	I180F (ATC-TTC)	RSc3268	Hypothetical protein	B41
CoB17	259709	C-G	Intergenic (+77/-149)	RSp0210 + / + RSp0211	Hypothetical protein/hypothetical protein	B41
CoB17	1565954	C-T	Intergenic (-112/+6)	RSp1230 - / - RSp1231	Hypothetical protein/putative transferase protein	B41
CoB17	1373289	G-A	Intergenic (-64/+19)	RSp1080 - / - RSp1081	Hypothetical protein/signal peptide protein	B41
CoB17	2825660	C-T	R48Q (CGG-CAG)	RSc2622	Integrase protein	B41
CoB17	1599495	C-T	S341L (TCG-TTG)	*nadB2*	L-aspartate oxidase (quinolinate synthetase B) oxidoreductase protein	B41
CoB17	2175967	C-T	L23L (CTC-CTT)	RSc1999	Lyase protein	B41
CoB17	2235814	G-A	S41L (TCG-TTG)	*nuoA*	NADH dehydrogenase subunit A	B41
CoB17	334658	C-T	F213F (TTC-TTT)	*nifD*	NifD nitrogenase molybdenum-iron protein alpha chain (Nitrogenase component I) (Dinitrogenase)	B41
CoB17	305491	G-A	L155L (CTC-CTT)	*nodA*	NodA N-acyltransferase	B41
CoB17	3629179	C-T	E257K (GAA-AAA)	RSc3369	Putative AVRPPHE avirulence protein	B41
CoB17	1832630	C-T	A297A (GCC-GCT)	RSp1457	Putative cation-efflux system signal peptide protein	B41
CoB17	587859	G-A	L641L (CTC-CTT)	RSp0474	Putative general secretion pathway GSPD-related protein	B41
CoB17	668997	T-A	Intergenic (+82/-71)	*icc* + / + RSp0535	Putative ICC protein/transcription regulator protein	B41
CoB17	1918596	A-T	D268K (GAT-AAA)	RSp1530	Putative L-ascorbate oxidase (ASCORBASE) oxidoreductase protein	B41
CoB17	1436807	C-T	A74A (GCG-GCA)	RSp1137	Putative RHS-related transmembrane protein	B41
CoB17	1808316	G-T	P254H (CCC-CAC)	*cysE1*	Serine acetyltransferase protein	B41

CoB17	1129391	G-A	D434N (GAT-AAT)	RSc1074	Signal peptide protein	B41
CoB17	241428	T-C	Intergenic (-229/-154)	pRALTA_0286 - / + pRALTA_0287	Transposase (fragment)/conserved hypothetical protein; putative Periplasmic sugar-binding domain protein	B41
CoB18	183055	C-T	S24F (TCT-TTT)	RSp0158	Hypothetical protein	B45, B46
CoB18	684983	G-A	R2Q (CGA-CAA)	RSp0542	Hypothetical protein	B45, B46
CoB18	1663440	C-T	P123L (CCT-CTT)	RSc1551	Hypothetical protein	B45, B46
CoB18	1797911	2 bp-TT	Coding (737-738/954 nt)	RSp1427	Hypothetical protein	B45, B46
CoB18	2386140	G-A	I258I (ATC-ATT)	RSc2204	Hypothetical protein	B45, B46
CoB18	2287799	C-T	E212K (GAG-AAG)	RSc2112	Metabolite transporter	B45, B46
CoB18	310043	A-G	Intergenic (-28/+385)	*nodC - / - nodB*	NodC N-acetylglucosaminyltransferase/NodB chitooligosaccharide deacetylase	B45, B46
CoB18	1638812	G-A	P533L (CCA-CTA)	RSp1288	Putative type-4 fimbrial biogenesis pily1-related protein	B45, B46
CoB18	611678	T-A	I48F (ATC-TTC)	*czcR*	Response regulator for cobalt zinc cadmium resistance transcription regulator protein	B45, B46
CoB20	544694	C-T	D24N (GAT-AAT)	RSp0430	Alcohol dehydrogenase	B45, B46
CoB20	1620019	G-A	E199K (GAG-AAG)	*fumC*	Fumarate hydratase	B45, B46
CoB20	2504339	G-A	F283F (TTC-TTT)	RSc2303	GSPD-like protein	B45, B46
CoB20	1690969	C-A	A383D (GCC-GAC)	*relA*	GTP pyrophosphokinase	B45, B46
CoB20	2846722	G-A	F388F (TTC-TTT)	*hutH*	Histidine ammonia-lyase	B45, B46
CoB20	1094068	C-T	L31L (CTC-CTT)	*hrcJ*	Hrp conserved lipoprotein HRCJ transmembrane	B45, B46
CoB20	248300	C-T	S281F (TCC-TTC)	RSc0225	Hypothetical protein	B45, B46
CoB20	253661	G-A	P290P (CCC-CCT)	RSp0205	Hypothetical protein	B45, B46
CoB20	2021703	2 bp-AA	Coding (588-589/900 nt)	*RSp1606*	Hypothetical protein	B45, B46
CoB20	2473508	G-A	E128K (GAG-AAG)	RSc2277	Hypothetical protein	B45, B46
CoB20	722435	G-A	Intergenic (-294/+149)	RSp0586 - / - RSp0587	Hypothetical protein/putative signal peptide protein	B45, B46
CoB20	1727668	G-A	E104K (GAG-AAG)	*nosY*	NosY transmembrane protein	B45, B46
CoB20	280495	G-A	S391L (TCA-TTA)	RSc0249	Putative calcium binding hemolysin protein	B45, B46
CoB20	939940	C-T	A253A (GCC-GCT)	RSp0749	Putative CLPA/B-type chaperone protein	B45, B46
CoB20	153903	C-T	E335K (GAG-AAG)	pRALTA_0179	Putative topoisomerase/recombinase/integrase	B45, B46
CoB20	1955536	G-A	Q716* (CAG-TAG)	*popF1*	Secreted protein POPF1	B45, B46
CoB20	85530	C-T	Pseudogene (160/1302 nt)	pRALTA_0090	Transposase fragment	B45, B46
CoM21	97654	C-T	P77S (CCA-TCA)	RSp0087	Hypothetical protein	M38, M50
CoM21	100357	C-T	L60L (CTG-TTG)	RSp0091	Hypothetical protein	M38, M50

CoM21	624178	C-T	G35R (GGG-AGG)	RSp0500	Hypothetical protein	M38, M50
CoM21	1173640	G-C	S1698C (TCC-TGC)	RSp0930	Hypothetical protein	M38, M50
CoM21	2222795	C-T	L454L (CTC-CTT)	*nuoM*	NADH dehydrogenase subunit M	M38, M50
CoM21	3434275	G-A	G139E (GGG-GAG)	RSc3175	Hypothetical protein	M38, M50
CoM24	230390	C-T	Coding (291-292/1227nt)	RSc0207	Putative transporter transmembrane protein	M38, M50
CoM24	393072	T-A	K239K (AAG-AAA)	RSp0296	Hypothetical protein	M38, M50
CoM24	571257	C-G	L199L (TTG-TTA)	RSp0457	Rhs-related protein	M38, M50
CoM24	584989	C-T	Coding (526-527/1044nt)	RSc0540	Serine/threonine protein kinase	M38, M50
CoM24	603178	C-T	Coding (440-443/507nt)	*pilA*	Type 4 fimbrial pilin signal peptide protein	M38, M50
CoM24	1210785	G-A	L27Q (CTG-CAG)	RSp0959	Anaerobic nitric oxide reductase transcription regulator	M38, M50
CoM24	1287580	2bp-AA	Intergenic (-232/+568)	*epsA* /RSp1021	EPS I polysaccharide export outer membrane transmembrane protein/hypothetical protein	M38, M50
CoM24	1431897	C-T	Intergenic (-339/+578)	*tISRso18* /RSp1137	ISRSO18-transposase protein/putative RHS-related transmembrane protein	M38, M50
CoM24	1716869	2bp-TT	F771F (TTC-TTT)	*kgd*	Alpha-ketoglutarate decarboxylase	M38, M50
CoM24	2966111	+4bp	G106E (GGA-GAA)	RSc2754	Hypothetical protein	M38, M50

[a] +, insertion of a nucleotide. bp, base pair. [b]up[x], intergenic mutations located x nucleotides upstream the gene indicated.

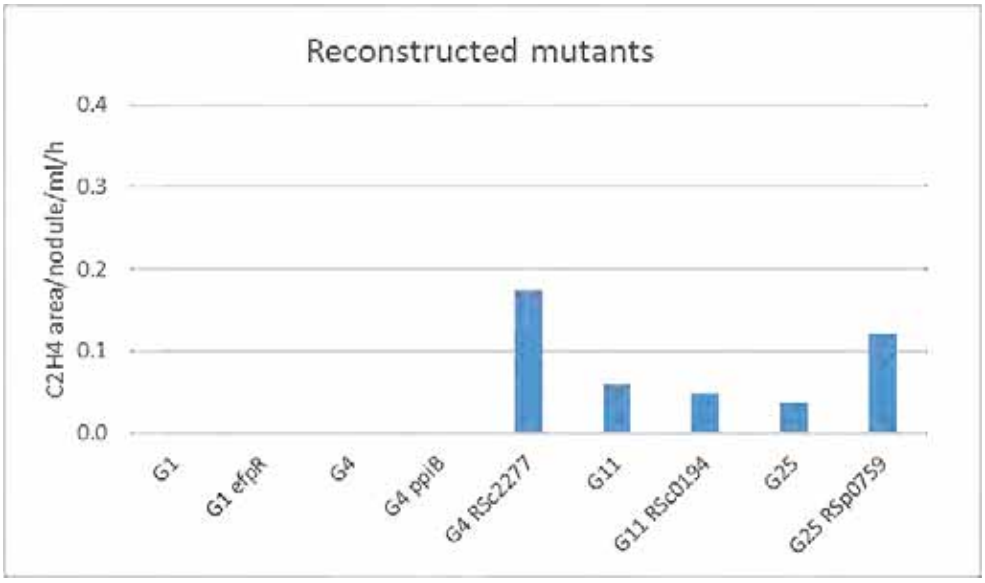


Figure 4.3. **Nitrogenase activity of the evolved clones carrying individually adaptive mutation improving within-host proliferation**. Data represent the reduction of acetylene per nodule at 15 days post inoculation. Six plants were exposed to acetylene during 4 hours, then pooled. Ethylene amounts were measured through gas chromatography and mass spectrometry. Data were obtained from 1 experiment.

and to a lesser extend RSp0759-P92S were able to increase the nitrogenase activity of bacteria in nodules.

For the lineages B and M, the main fixation shifts occurred after cycle 35 and no sequencing data were available for the lineage X. Using the Illumina sequencing technology, we sequenced whole-populations every two cycles between cycles 35 and 47 for the B and M lineages and between cycles 1 and 5 for the X lineage (figure 4.4). As in previous work (Doin de Moura et al., 2022), we found many mutations keeping low or intermediate frequencies in the evolved populations. To analyze mutational trajectories, we focused on mutations having a frequency of at least 30% frequency in at least one evolution cycle in evolved populations. The molecular evolution of these three lineages followed a pattern previously described during the first cycles of the evolution (Doin de Moura et al., 2022) with a constant accumulation of mutations along the evolution cycles and the formation of mutational cohorts, *i.e.* groups of mutations approximatively following the same frequency trajectories over time. The occurrence of several sweeps between the cycles 35 and 47 for the lineages B and M suggests that selection continues to act in these populations (figures 4.4a and 4.4b, respectively). We identified 5 cohorts in lineage X, although they are not as well supported as in other lineages since we analyzed only 3 cycles (figure 4.4c).

In order to investigate the genetic bases of nitrogenase activity, we searched for the presence of the identified mutational cohorts in the genome sequence of evolved clones surrounding nitrogen fixation shifts (Tables 4.2 to 4.5) and we constructed phylogenetic trees for each lineage (figures 4.5 to 4.7). Comparing the presence/absence of the different mutational cohorts in individual clones to their nitrogenase activity allowed us to generate a list of candidate mutations potentially involved in this phenotype.

In the lineage B, highest levels of nitrogenase activity were observed for the clones B41, B45 and B46. The phylogenetic tree of the B evolved clones indicated that the clones B45 and B46 did not derive from the same subpopulation than the clone B41 (figure 4.5). The clone B41 carries the cohort CoB17 (having 28 mutations) while the clones B45 and B46 (Figure 4.4a, figure 4.5, table4.1, table 4.2) share the cohorts CoB18 and CoB20 (containing 9 and 17 mutations, respectively). It is noteworthy that, the CoB18/CoB20 and CoB17 co-exist in populations from cycles 31 to 47. At cycle 47, these two competing sub-populations each reached almost 50% of frequency in the evolved population, with CoB18/CoB20 having a

Table 4.2. Mutational cohorts present in the evolved clones from the lineage B

Clone	CoB01	CoB02	CoB03	CoB04	CoB05	CoB06	CoB07	CoB08	CoB09	CoB10	CoB11	CoB12	CoB13	CoB14	CoB15	CoB16	CoB17	CoB18	CoB19	CoB20	CoB21	CoB22	CoB23	CoB24
B01	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B02	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B03	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B04	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B05	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B06	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B07	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B08	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B09	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B10	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B11	1	1	1	-	1	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B12	1	1	1	-	1	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B13	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B14	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B15	1	1	1	-	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B16	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B17	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B18	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B19	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B20	1	1	1	-	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B21	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B22	1	1	1	-	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B23	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B24	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B25	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B26	1	1	1	-	1	1	-	1	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
B27	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B28	1	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B29	1	1	1	-	1	1	1	-	-	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-
B30	1	1	1	-	1	1	1	-	-	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-
B31	1	1	1	-	1	1	1	-	-	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-
B32	1	1	1	-	1	1	-	1	-	-	1	-	-	-	1	-	-	-	-	-	-	-	-	-
B33	1	1	1	-	1	1	-	1	-	-	1	-	-	-	1	-	-	-	1	-	-	-	-	-
B35	1	1	1	-	1	1	1	-	-	1	-	-	1	-	-	1	-	-	1	-	-	-	-	-
B40	1	1	1	-	1	1	1	-	-	1	-	-	1	-	-	1	-	-	-	-	-	-	-	-
B41	1	1	1	-	1	1	1	-	-	1	-	1	-	1	-	-	1	-	-	-	-	-	-	-
B45	1	1	1	-	1	1	1	-	-	1	-	-	1	-	-	1	-	1	-	1	-	-	1	-
B46	1	1	1	-	1	1	1	-	-	1	-	-	1	-	-	1	-	1	-	1	-	1	-	-
B50	1	1	1	-	1	1	1	-	-	1	-	-	1	-	-	1	-	1	-	1	-	-	-	1

Table 4.3. Mutational cohorts present in the evolved clones from the lineage G

Clone	CoG1	CoG2	CoG3	CoG4	CoG5	CoG6	CoG7	CoG8	CoG9	CoG10	CoG11	CoG12	CoG13	CoG14	CoG15	CoG16	CoG17	CoG18	CoG19	CoG20	CoG21
G1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G2	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G3	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G4	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G5	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G6	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G7	1	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G8	1	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G9	1	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G10	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G11	1	1	-	1	-	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G12	1	1	-	1	-	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G13	1	1	-	1	-	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G14	1	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G15	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G16	1	1	-	1	-	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
G17	1	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G18	1	1	-	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-
G19	1	1	-	1	-	1	1	1	-	-	1	-	1	1	-	-	-	-	-	-	-
G20	1	1	-	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-
G21	1	1	-	1	-	1	1	1	-	-	1	-	1	1	-	-	-	-	-	-	-
G22	1	1	-	1	-	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
G23	1	1	-	1	-	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
G24	1	1	-	1	-	1	1	1	-	-	1	-	1	1	-	-	-	-	-	-	-
G25	1	1	-	1	-	1	1	1	-	-	1	-	1	1	-	-	-	-	-	-	-
G26	1	1	-	1	-	1	1	1	-	-	1	-	1	1	-	-	-	-	-	-	-
G27	1	1	-	1	-	1	1	1	-	-	1	-	1	1	-	-	-	-	-	-	-
G28	1	1	-	1	-	1	1	1	-	-	1	-	1	-	1	-	1	-	1	-	-
G29	1	1	-	1	-	1	1	1	-	-	1	-	1	-	1	-	1	-	1	-	-
G30	1	1	-	1	-	1	1	1	-	-	1	-	1	-	1	-	1	-	1	-	-
G31	1	1	-	1	-	1	1	1	-	-	1	-	1	1	-	-	-	-	-	-	-
G32	1	1	-	1	-	1	1	1	-	-	1	-	1	-	1	1	-	1	-	-	-
G33	1	1	-	1	-	1	1	1	-	-	1	-	1	-	1	1	-	-	-	-	-
G35	1	1	-	1	-	1	1	1	-	-	1	-	1	-	1	1	-	1	-	-	-

Table 4.4. Mutational cohorts present in the evolved clones from the lineage M

Clone	CoM 1	CoM 2	CoM 3	CoM 4	CoM 5	CoM 6	CoM 7	CoM 8	CoM 9	CoM 10	CoM 11	CoM 12	CoM 13	CoM 14	CoM 15	CoM 16	CoM 17	CoM 18	CoM 19	CoM 20	CoM 21	CoM 22	CoM 23	CoM 24	CoM 25	CoM 26	CoM 27	CoM 28
M2	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M4	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M6	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M8	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M10	1	1	1	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M12	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M14	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M16	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M35	1	1	1	1	-	-	1	-	-	1	-	-	-	-	-	1	1	1	1	-	-	-	-	-	-	-	-	-
M36	1	1	1	1	-	-	1	-	-	1	-	-	1	-	-	1	1	1	1	-	-	1	1	-	-	-	-	-
M38	1	1	1	1	-	-	1	-	-	1	-	-	1	-	-	1	1	1	-	-	1	-	-	1	-	-	-	-
M50	1	1	1	1	-	-	1	-	-	1	-	-	1	-	-	1	1	1	-	-	1	-	-	1	-	1	1	-

Table 4.5. Mutational cohorts present in the evolved clones from the lineage X

Clone	CoX01	CoX02	CoX03	CoX04
X03	-	-	-	-
X06	-	1	-	1

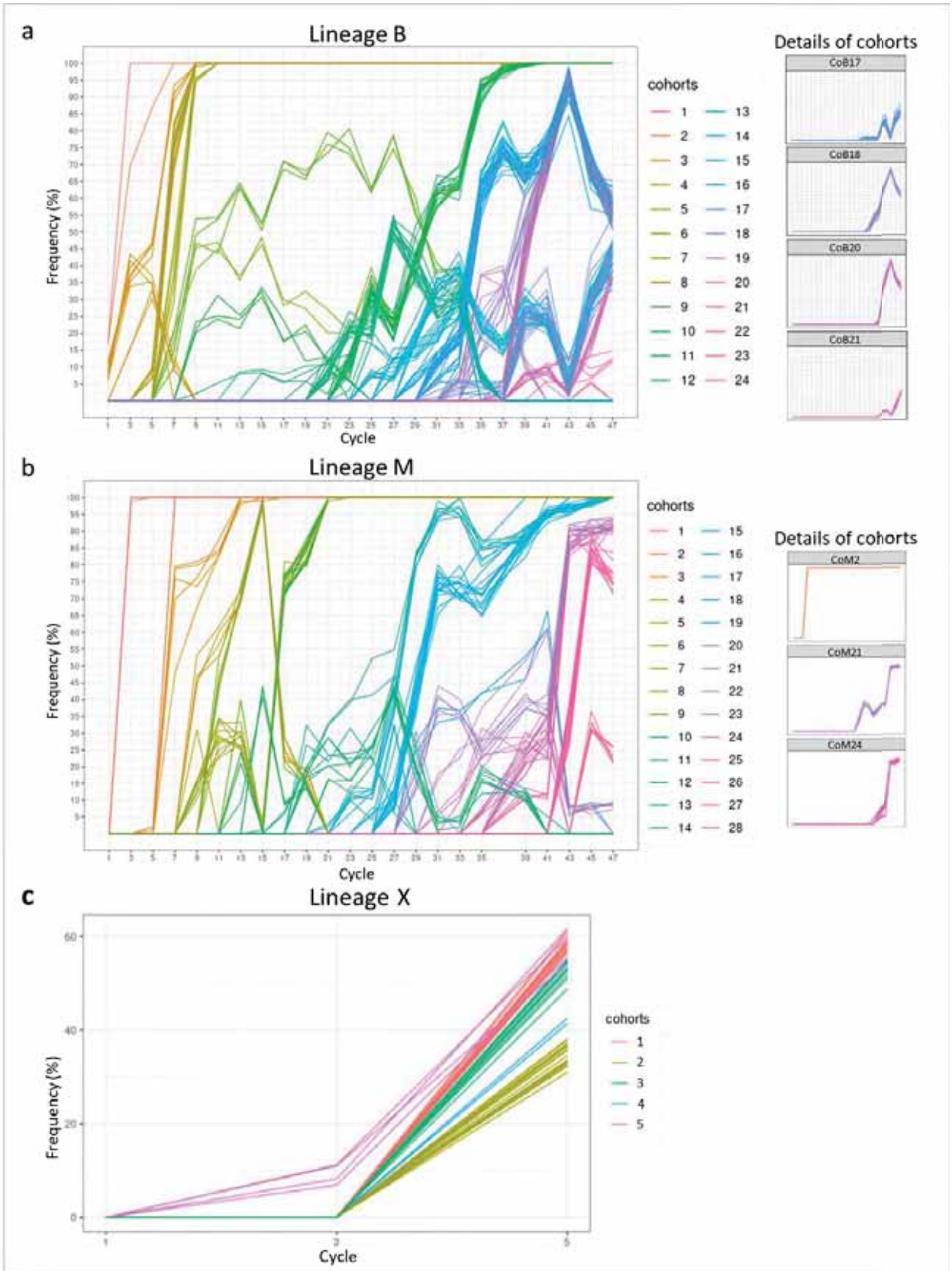


Figure 4.4 : Dynamics of molecular evolution in the evolved lineages B and M until the cycle 47 and in the lineage X between the cycles 1 and 5. Allele frequency trajectories of mutations that attained a frequency of 30% in at least one population of the B, M and X lineages (**a**,**b**,**c**, respectively). Mutations with similar trajectories were clustered in cohorts and represented by different colors. For simplicity, mutations travelling alone were also called cohorts. Details of the cohorts cited in the text are represented on the right side of the figure.

downward trajectory and CoB17 (together with the newly appeared cohort CoB21) an upward trajectory. The clone B50, which is not fixing nitrogen, also carries the cohort CoB18/CoB20. However, this clone possesses a specific frameshift mutation in the regulator of the *nif-fix* genes *nifA*, which may be responsible for its nitrogen fixation defect.

In the lineage M, we searched for conserved cohorts between the nitrogen-fixing clones M38 and M50 but absent in the non-fixing clone M36. The cohorts CoM21 and CoM24 (containing 6 and 10 mutations, respectively) met these criteria (figure 4.4.b, figure 4.7, table 4.1, table 4.4). Besides that, the clone X3, which displayed a first level of nitrogenase activity, carries no mutational cohorts identified in this lineage. However, it carries 10 specific mutations compared to its nodulating ancestor. An increased level of nitrogen fixation was observed in the clone X6 (figure 4.2f), which carries the cohorts CoX02 and CoX04 (containing 18 and 2 mutations, respectively) in addition to 5 mutations specific to this clone (Figure 4.4c, table 4.5).

Interestingly, a striking mutational convergence in the gene RSc2277 was observed in almost all clones showing a shift in nitrogenase activity, namely the clones B41 (cohort CoB17), B45 and B46 (cohort CoB20), G25 (cohort CoG4) and X6 (specific mutation) (figures 4.5 to 4.7, table 6). In lineage G, an adaptive mutation in the gene RSc2277 arose early in the experiment (CoG4, Doin de Moura et al., 2022) and all clones tested beyond G25 showed some level of nitrogen fixation (figure 4.2c). Therefore, even if all nitrogen-fixing clones carry the RSc2277-V321G mutation, we do not have enough ARA data to precisely locate the nitrogen fixation shift. Interestingly, another mutation in RSc2277, although not present in the clone X6, is present at high frequency (59%) in the population X after 5 cycles and belongs to the cohort CoX01 (Figure 4.4c and table 4.6). In the lineage M, a mutation in RSc2277 arose early in CoM2 but was not associated with a nitrogen fixation shift (Figure 4.7). These findings led us to further investigate the role of RSc2277 in the increase in nitrogenase activity.

The effect of the RSc2277-V321G mutation on nitrogenase activity is dependent on the *efpR*-E66K mutation

Interestingly, in the evolved lineages B, G and X, mutations in RSc2277 appeared after adaptive mutations affecting the EfpR pathway, either after the mutation *efpR*-E66K (lineages G and X) or after the mutation up[115]-RSc0965 (lineage B), both improving the intracellular infection of

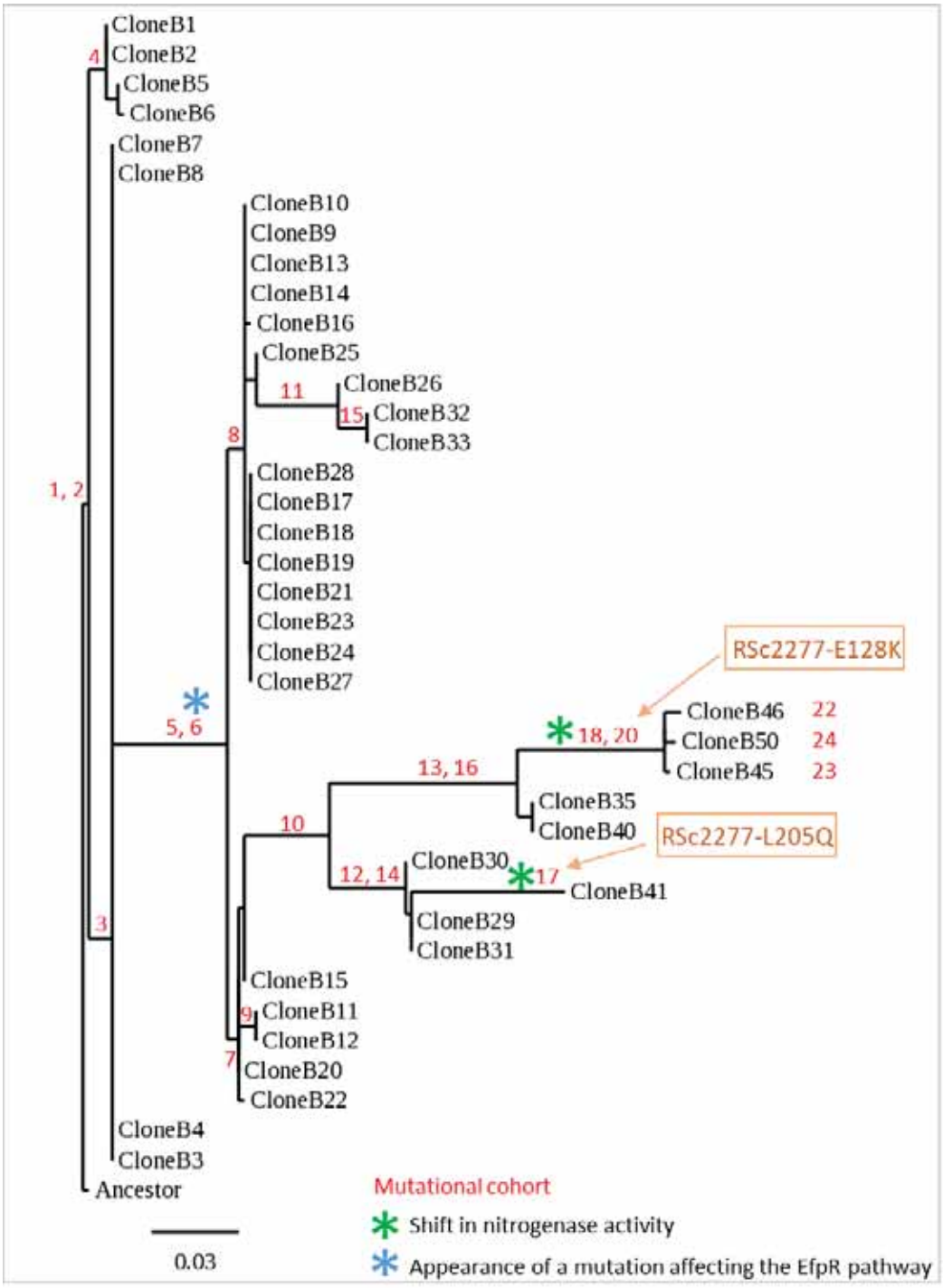


Figure 4.5. **Phylogenetic tree of evolved clones from the lineages B during the evolution experiment**. Evolved clones are indicated in black. Green stars indicate shifts in nitrogenase activity. The blue star indicates the appearance of the mutation up[115]-RSc0965, affecting the EfpR pathway. Numbers in red indicate the cohorts present in each clone. Mutations in the gene RSc2277 are represented in orange and its presence in the cohorts is indicated by the orange arrows. The black bar represents the scale of genetic variations using the maximum likelihood method applied on artificial sequences constructed by concatenating all mutated allele positions appearing in the experiment.

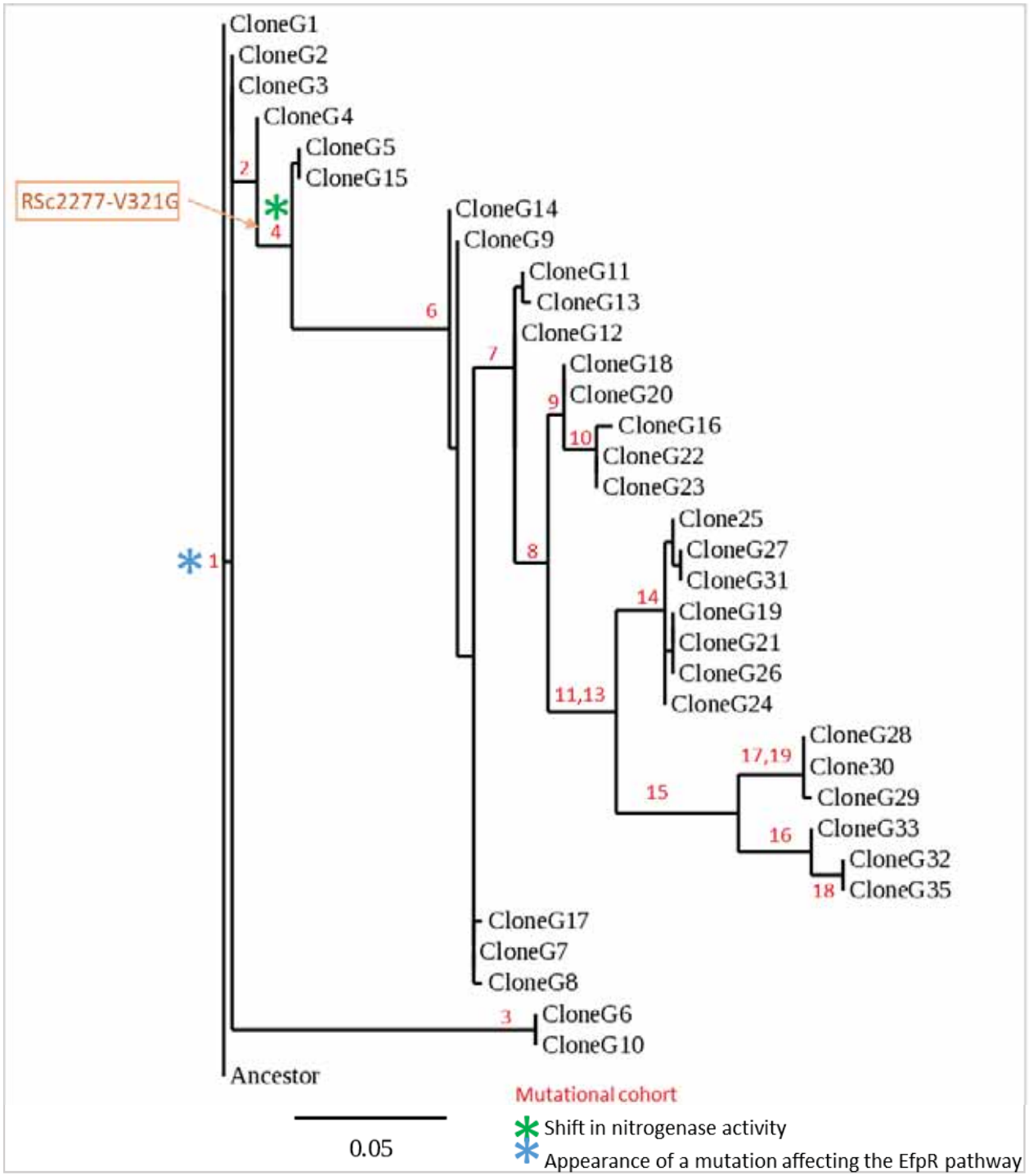


Figure 4.6. **Phylogenetic tree of evolved clones from the lineages G during the evolution experiment.** Evolved clones are indicated in black. Numbers in red indicate the cohorts present in each clone. Green stars indicate shifts in nitrogenase activity. The blue star indicates the appearance of the mutation *efpR*-E66K, affecting the EfpR pathway. Mutations in the gene RSc2277 are represented in orange and its presence in the cohorts is indicated by the orange arrows. The black bar represents the scale of genetic variations using the maximum likelihood method applied on artificial sequences constructed by concatenating all mutated allele positions appearing in the experiment.

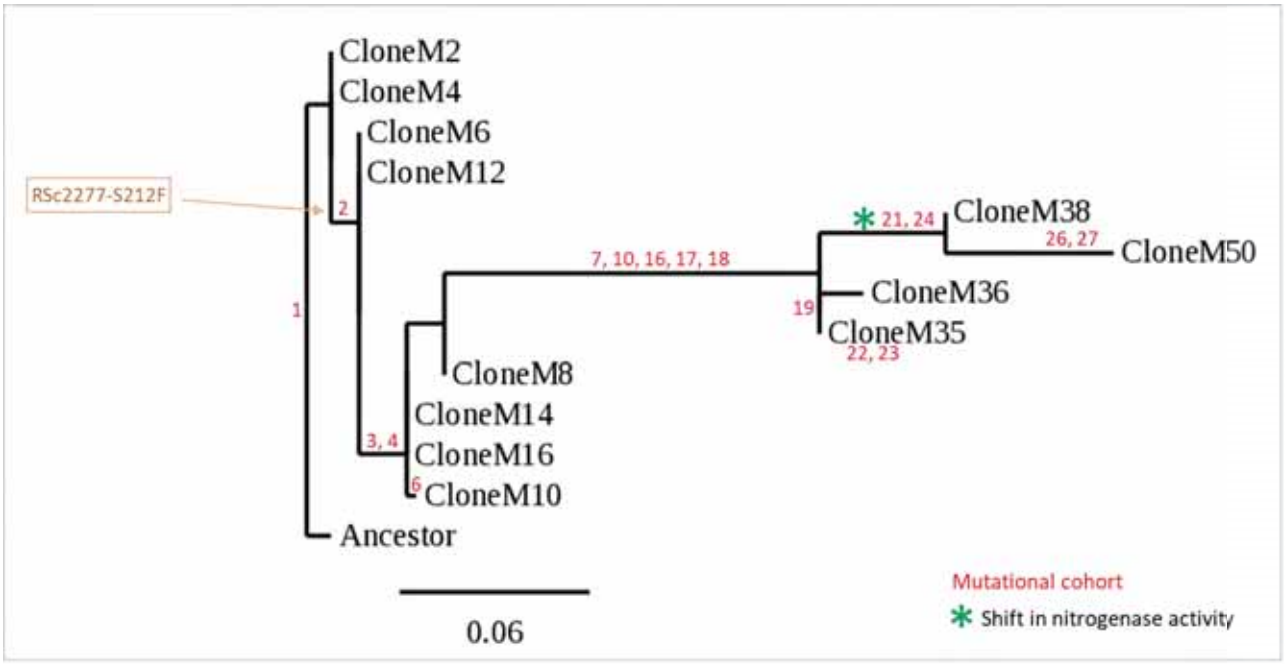


Figure 4.7. **Phylogenetic tree of evolved clones from the lineages M during the evolution experiment**. Evolved clones are indicated in black. Green stars indicate shifts in nitrogenase activity. Numbers in red indicate the cohorts present in each clone. Mutations in the gene RSc2277 are represented in orange and its presence in the cohorts is indicated by the orange arrows. The black bar represents the scale of genetic variations using the maximum likelihood method applied on artificial sequences constructed with the mutated allele positions appearing in the experiment.

nodules (Capela et al., 2017). In the lineage G, the mutation V321G on the gene RSc2277 generates another shift in the infectivity of the evolved clones (Doin de Moura et al., 2022). This prompted us to hypothesize that the previously observed shift in nitrogenase activity generated by the RSc2277-V321G mutation in the lineage G is linked to an additional increase in the number of bacteria per nodule, leading the amount of reduced acetylene to detectable levels, and thus could be dependent on the first shift in infectivity generated by the mutations affecting the EfpR pathway.

To test this hypothesis, we introduced the RSc2277-V321G mutation individually or combined with the *efpR*-E66K mutation in a GMI1000 pRalta *hrpG* background and compared their relative *in planta* fitness, their effect in infectivity and acetylene reduction ability. The mutation RSc2277-V321G increased the *in planta* fitness of both GMI1000 pRalta *hrpG* and GMI1000 pRalta *hrpG efpR*-E66K strains (Figure 4.8), indicating that this mutation is adaptive for symbiosis in both backgrounds. Contrary to this observation, the mutation RSc2277-V321G improved the nitrogenase activity in the GMI1000 pRalta *hrpG-efpR*-E66K background but only to very weak levels close to zero in the GMI1000 pRalta *hrpG* background (figure 4.9). Interestingly, preliminary infectivity data are in accordance with this observation, indicating that the mutation RSc2277-V321G increased infectivity at 15 dpi in both GMI1000 pRalta *hrpG efpR*-E66K and GMI1000 pRalta *hrpG* strains, but has an additive effect on infectivity in the efpR-E66K background. This showed that the combination of the mutation RSc2277-V321G and *efpR*-E66K mutation allowed to each higher levels of infectivity (Figure 4.10). It is possible that this increase in infectivity generated by the mutation RSc2277-V321G in an *efpR* background, lead to higher amounts of nitrogenase activity in the nodules formed the GMI1000 pRalta *hrpG efpR*-E66K RSc2277-V321G strain.

Discussion

Mutualistic nitrogen-fixing symbiosis is a multistep process, mediated by several genetic determinants of both plant and bacterial partners. Initiation of rhizobium-legume symbiosis was deeply studied in a number of model systems and in most cases is principally mediated by the production of bacterial Nod factors specifically recognized by plant receptors (Arrighi et al., 2006; Limpens et al., 2003; Madsen et al., 2003; Moling et al., 2014; Radutoiu et al., 2003; Smit et al., 2007). By contrast, late stages of symbiosis were less well described and

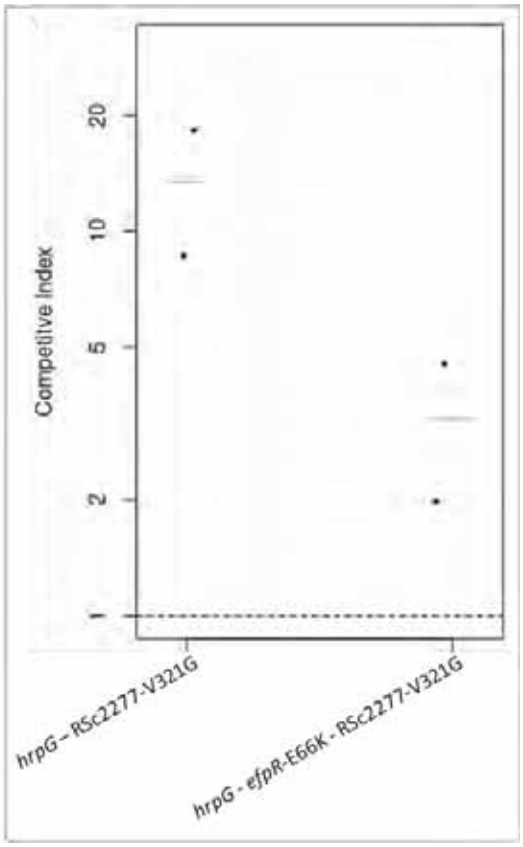


Figure 4.8. **Relative *in planta* fitness of reconstructed strains carrying the mutation RSc2277-V321G**. The tested mutations was reconstructed on both the *Ralstonia solanacearum* GMI1000 pRalta *hrpG* and GMI1000 pRalta *hrpG efpR*-E66K backgrounds. *Mimosa pudica* plants were co-inoculated with the reconstructed strains and respective parental strains. Then, nodules were harvested at 21 days post inoculation, then pooled and crushed together. Finally, nodule crushes were plated, allowing the measurement of the proportion of each strain in the final nodule bacterial population. Preliminary data obtained from 2 independent experiments.

probably less conserved between legumes, some underlying mechanisms being host-plant dependent (Burghardt et al., 2017; diCenzo et al., 2015; Montiel et al., 2017). In our evolution experiment, these late symbiotic steps are not fully carried out by our evolved clones, even after 60 nodulation cycles, representing *ca*. 1500 bacterial generations (Remigi et al., 2014). Acetylene reduction assays showed that the nitrogenase activity was not null in four lineages but was extremely low and transient. A possible reason that may explain this nitrogen fixation defect is that bacteria only survive for a (very) short time in nodule cells resulting in a transient metabolic activity and thus transient nitrogenase activity. Such a low persistence of *Ralstonia* evolved clones could be due to the absence of some important late symbiotic genes in *R. solanacearum*. The rhizobial genetic determinants involved in long-term intracellular persistence were relatively poorly described. They may include genes needed to survive under high osmotic pressure (Fougère and Le Rudulier, 1990; Ohwada et al., 1998; Tian et al., 2022) and physical pressure due to the cell wall and outward-directed turgor pressure (Parniske, 2018) and microoxic conditions, which prevail in nodules (Soupene et al., 1995). In nodules, aerobic respiration of bacteria is ensured by plant hemoproteins called leghemoglobins (Dakora, 1995; Ott et al., 2009; Ott et al., 2005), which deliver oxygen to bacteroids, and by bacterial cytochrome oxidases with very high affinity for oxygen (Cosseau & Batut, 2004; Preisig et al., 1996; Torres et al., 2013; Zufferey et al., 1996). The genome of *R. solanacearum* carries one operon encoding a cytochrome c oxidase of cbb3-type. However, its affinity for oxygen compared to the *C. taiwanensis* CcoNOQP cytochrome c oxidase is not known. Respiration of bacteria in nodules is essential to generate high amounts of ATP necessary for reducing atmospheric nitrogen into ammonia. Production of energy also relies on the capacity of bacteria to use nitrogen and carbon sources supplied by the plant (Ding et al., 2012; Dymov et al., 2004; Fagorzi et al., 2020; R. T. Green et al., 2019; J. I. Jiménez-Zurdo et al., 1995; Köhler et al., 2000; Liu et al., 2017; Prell and Poole, 2006; 1981; Schulte et al., 2021; Schulte et al., 2022; Wielbo et al., 2007). Moreover, in some legumes, such as *Medicago* and *Aeschynomene*, it was demonstrated that nodule bacteria have to survive to antimicrobial peptides called NCR (Nodule-specific Cysteine Rich) peptides, specifically produced in nodules and triggering terminal bacteroid differentiation (Czernic et al., 2015; Guefrachi et al., 2015; Van de Velde et al., 2010). In these cases, only compatible symbionts possessing a specific transporter (BacA or BclA) can survive within nodule cells and differentiate into nitrogen-fixing bacteroids

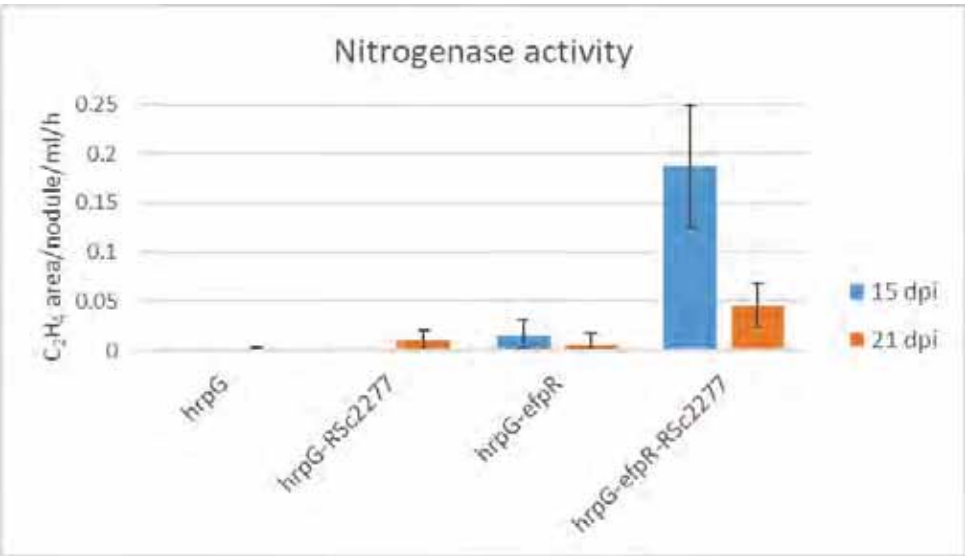


Figure 4.9. **Nitrogenase activity of the reconstructed strains GMI1000 pRalta *hrpG* (CBM1627), GMI1000 pRalta *hrpG* RSc2277-V321G (RCM3467), GMI1000 pRalta *hrpG efpR*-E66K (RCM1865) and GMI1000 pRalta *hrpG efpR*-E66K RSc2277 –V321G (RCM3475)** at 15 (blue bars) and 21 days post inoculation (orange bars). Six plants were exposed to acetylene during 4 hours, then pooled. Ethylene amounts were measured through gas chromatography and mass spectrometry. Preliminary data were obtained from 2 independent experiments.

(Ardissone et al., 2011; Guefrachi et al., 2015; Karunakaran et al., 2010). Although such NCR peptides were not found on the *Mimosa pudica* genome (unpublished data) and *C. taiwanensis* has no BacA-like transporter, we cannot rule out that other peptides may be produced in *Mimosa* nodules to control bacterial proliferation and that specific bacterial genes may be required to survive to these peptides (Keller and Libourel et al., 2022).

Mutational convergence in evolution experiments is often a sign that mutations are adaptive, and RSc2277 is one of the genes showing the highest level of convergence in our experiment (Doin de Moura et al. 2022). Mutations in the RSc2277 gene were convergently associated with nitrogenase activity improvements. Although this association was not always strict (*e.g*. some clones from lineage M have a mutation in this gene without having detectable ARA activity) we showed that one of these mutations (V321G, from lineage G) is sufficient to confer a low level of nitrogen fixation to a clone already carrying the *efpR* E66K mutation. However, the role of mutations in this gene in the B, M and X lineages and their effect on infectivity and nitrogen fixation remains to be demonstrated. Interestingly, in addition to one deletion of 1 bp in the clone X06, all mutations appearing in this gene are non-synonymous, potentially causing the loss of function of the gene. This hypothesis remains to be tested by comparing the mutants carrying point mutation with the RSc2277 deletion mutant. Initially, very little information about this gene encoding a putative transporter was available. In the chapter 3 of this thesis, we have started the functional characterization of this gene and showed that the RSc2277-V321G mutation has pleiotropic effects, increasing bacterial swimming and swarming motility, decreasing exopolysaccharide production, enhancing the metabolism of some nitrogen and carbon compounds and allowing a better survival under acidic conditions. However, nothing is known about the functioning of this protein, its regulation or even its links with the expression of other genes. In the *R. solanacearum* genome, this gene is located near the *rag* genes, encoding proteins putatively involved in polysaccharide synthesis. Homologs of *rag* genes in *Pseudomonas aeruginosa*, the *pel* genes, have been linked to the biosynthesis of a novel glucose-rich polysaccharide component involved in the formation of resistant biofilms (Friedman and Kolter, 2004; Vasseur et al., 2005). Whether in *R. solanacearum* RSc2277 is linked to the *rag* genes and the transport of some polysaccharides and biofilm formation or whether this transporter is involved in global regulatory mechanisms causing pleiotropic effects remains to be tested.

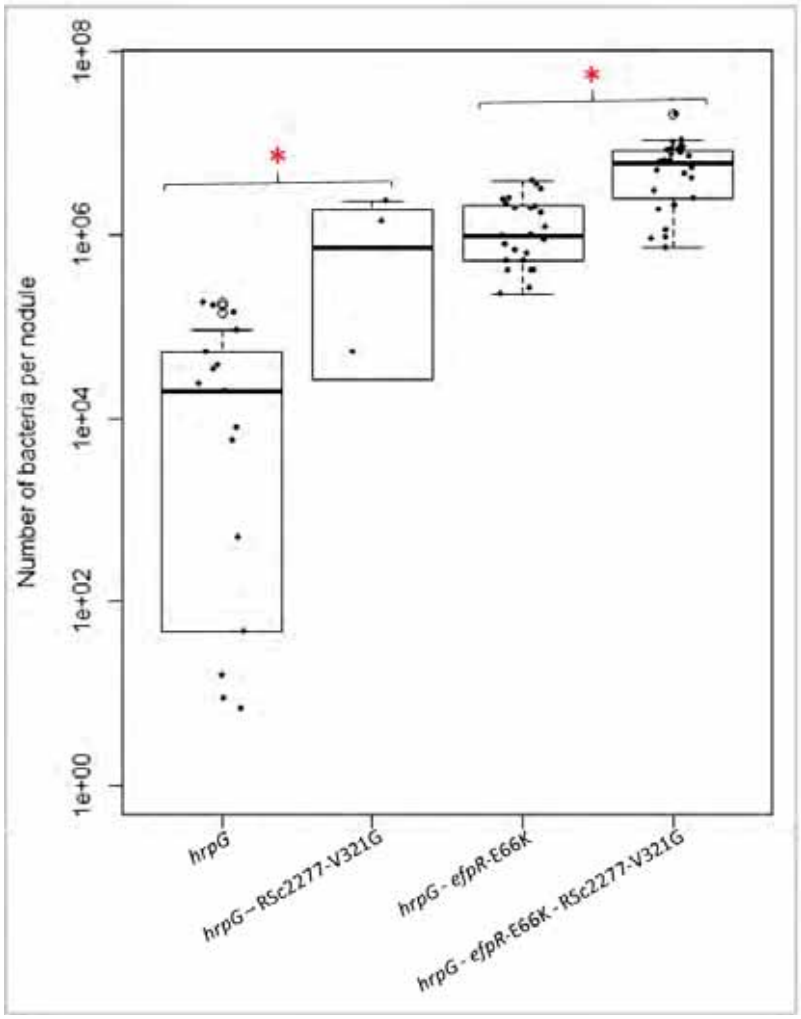


Figure 4.10. **Infectivity (number of bacteria per nodule) of reconstructed strains carrying mutations improving the infectivity of evolved clones in the lineage G**. All tested mutations were reconstructed in the *Ralstonia solanacearum* chimeric strain GMI1000 pRalta. *Mimosa pudica* plants were singly inoculated with the reconstructed strains. Red stars indicate strains showing improvements in infectivity compared to their parental strain (Wilcoxon test, P< 0.05). Infectivity data were from 2 independent experiments except for the mutant *hrpG*-RSc2277-V321G which was measured in only one experiment and need to be repeated. Infectivity data were produced in collaboration with A.C. Cazalé in the team

Nitrogen-fixing symbiosis is a complex process, and thus we can expect interactions between adaptive mutations to occur in this evolution experiment. For example, the realization of some symbiotic steps allowed by a first adaptive mutation may be a pre-requisite for the effect of another adaptive mutation. Surprisingly, the highly adaptive mutation RSc2277-V321G does not seem to be dependent on the *efpR*-E66K mutation to improve the *in planta* fitness and infectivity of evolved bacteria, but has an additive effect to the *efpR* mutationon the infectivity of the evolved clones. This additive effect which led to a higher number of bacteria per nodule may then allow the detection of nitrogenase activity. Alternatively, a functional link between the two genes may explain the detection of nitrogenase activity. Indeed, since both *efpR* and RSc2277 mutations confer metabolic gains in *Ralstonia* (Capela et al., 2017; and chapitre II of this manuscript), these mutations may act in synergy to increase the metabolic activity of evolved clones, which may be responsible for the observed levels of nitrogen fixation.

In conclusion, despite the rapid improvement of early symbiotic traits (nodulation and intracellular infection of nodules), nitrogen fixation was not achieved in the experiment or achieved in very low levels, even after several hundreds of bacterial generations under plant selection pressure. Since nodulation competitiveness is strongly selected at the entry of the roots in the nitrogen-fixing symbiosis and nitrogen fixation is not selected at the entry point (Daubech et al., 2017), we can hypothesize that this symbiotic trait is under very weak selection in our system, with nitrogen-fixing clones being outcompeted by the most competitive strains for nodulation. Alternatively, important genes from *Cupriavidus taiwanensis*, in addition to the essential symbiotic ones already present in the symbiotic plasmid, may be absolutely required to achieve the late symbiotic stages. In this case, complementation of evolved bacteria with specific *C. taiwanensis* candidate genes (*ccoNOQP* or other genes to be identified from comparative genomics of natural *Mimosa* symbionts) or random DNA fragments could allow answering this question.

Material and methods

Bacterial culture and growth

Strains and plasmids used in this study are listed in Sup Table 4.1. *Ralstonia solanacearum* strains were grown at 28°C either on BG medium (Boucher et al., 1985) or on synthetic MP medium (Plener et al., 2010) supplemented with 2% glycerol (for bacterial transformation).

Table 4.6. Mutations detected in the gene RSc2277 in the evolved lineages

Lineage	Position	Mutation[a]	Change effect	Cohort	Clone
B	2473740	T-A	L205Q (CTG-CAG)	CoB17	B41
B	2473508	G-A	E128K (GAG-AAG)	CoB20	B45, B46, B50
G	2474343	T-C	L406P (CTG/CCG)	CoG3	G6, G10
G	2474088	T-G	V321G (GTG-GGG)	CoG4	G5 →G60
G	2473174	C-T	F16F (TTC-TTT)	CoG8	G16→G60
K	2473247	C-T	R41C (CGC-TGC)	-	K3
K	2473998	C-T	P291L (CCC-CTC)	CoK1	K4→K60
M	2473761	C-T	S212F (TCC-TTC)	CoM2	M6→M60
X	2473453	del1 bp	Frameshift pos. 327[b]	-	X6
X	2474039	G-T	E305*(GAG-TAG)	CoX1	-

[a] del, deletion of one nucleotide. [b]nucleotides after the start codon ATG.

Antibiotics were used at the following concentrations: trimethoprim at 100 µg/ml and kanamycin at 50 µg/ml.

Evolution experiment and cycles

Additional evolution cycles were performed as described by (Marchetti et al., 2014). Briefly, in each cycle, bacterial populations from cycle 35 were thawed from -80°C freezer and revived on BG plates before inoculating them on *Mimosa pudica* plants. 5.10^5 bacteria per plants were inoculated on 30 plants for each lineage. Nodules formed were harvested 21 days after inoculation, surface sterilized with 2.6% sodium hypochlorite during 15 min and crushed in 1 ml of sterile water. 1/10 of the nodule crush was re-inoculated on a new set of 30 plants. Additionally, serial dilutions of the nodule crushes were plated, allowing the selection and isolation of one clone representative of each population. The rest of nodule crushes and a 24h-culture of the nodule crushes were stored at -80°C. Using this protocol, bacterial populations of lineages B, F, G, K and M were evolved until cycle 60. Besides that, another lineage, the lineage X, was started using the strain RCM1865 (GMI1000 pRalta *hrpG efpR*-E66K) as ancestor and evolved for 15 nodulation cycles.

Sequencing of bacterial populations and clones and identification of mutations

To sequence the bacterial populations B37, B39, B42, B43, B45, B47, M37, M39, M41, M43, M45, M47, X1, X3 and X5 and the evolved clones B40, B41, B45, B46, B50, M36, M38, M50, X3 and X6, nodule bacterial populations and clones were grown overnight in BG medium supplemented with trimethoprim. Then, bacterial DNAs were extracted from 1 mL of culture using the Wizard Genomic DNA Purification Kit (Promega) and sequenced at the GeT-PlaGe core facility (https://get.genotoul.fr/), INRAE Toulouse. DNA-seq libraries have been prepared according to Illumina's protocols using the Illumina TruSeq Nano DNA HT Library Prep Kit. Briefly, DNA was fragmented by sonication, size selection was performed using SPB beads (kit beads) and adapters were ligated to be sequenced. Library quality was assessed using an Advanced Analytical Fragment Analyzer and libraries were quantified by qPCR using the Kapa Library Quantification Kit (Roche). Sequencing has been performed on a NovaSeq6000 S4 lane (Illumina, California, USA) using a paired-end read length of 2x150 pb with the Illumina NovaSeq Reagent Kits.

Identification of mutations were performed as described by (Doin de Moura et al., 2022) using Breseq v0.33.1 (Deatherage and Barrick, 2014) with default parameters, either using the polymorphism mode (for whole-population sequences) or the consensus mode (for individual clones). Mutation lists were curated manually in order to remove mutations present in the ancestral strains as well as false positive hits arising from reads misalignments in low complexity regions.

Phylogeny of evolved clones

Based on the mutations that appeared during the evolution experiment, sequences were reconstructed for each clone using either the mutated or the wild-type allele. Artificial sequences were constructed taking only the mutations appearing in the experiment and giving to each clone the wild-type or the mutant allele based on their genome sequences. Phylogenetic trees were reconstructed online (www.phylogeny.fr) using the maximum likelihood (ML) heuristic search under the model LG (Le & Gascuel, 2008) with C-rate variation among sites (Yang, 1994), as implemented in PhyMLv3.0 (Guindon et al., 2010).

Construction of mutants

The mutation V321G in the gene RSc2277 was reconstructed in the strains GMI1000 pRalta *hrpG* (CBM1627) and GMI1000 pRalta *hrpG efpR*-E66K (RCM1865) using the MuGent method (Dalia et al., 2014) as described in previous work (Doin de Moura et al., 2022). Briefly, the mutated region and constitutively expressed reporter genes (GFP or mCherry) were introduced by natural co-transformation of *R. solanacearum* strains. The presence of the RSc2277-V321G mutation was selected by PCR screening and verified by Sanger Sequencing. Oligonucleotides used for these constructions are described in (Doin de Moura et al., 2022).

Relative *in planta* fitness assays

Relative *in planta* fitness assays were performed as described by (Doin de Moura et al., 2022). Briefly, *Mimosa pudica* plantlets cultivated in tubes containing 20 mL of solid Fahreaus medium and 40 mL of liquid Jensen medium diluted 1:4 with sterile water were co-inoculated with the bacterial strains in equivalent proportion (ca. 5 x 10^5 bacteria of each strain per plant, in tubes containing 2 plantlets). Plants inoculated with both strains were incubated in a culture

chamber at 28°C under a photoperiod day/night of 16h/8 h. Twenty-one days after inoculation, nodules from 20 plants were pooled and crushed in sterile water, diluted and spread on selective solid medium using an easySpiral automatic plater (Interscience). Colonies from each strain were screened by fluorescence using a stereo-zoom microscope (AxioZoom – V16, Zeiss, Oberkochen, Germany).

Acetylene reduction assays

Acetylene reduction assays were performed as described by (Turner and Gibson, 1980). Briefly, plantlets of *Mimosa pudica* were inoculated individually with 5.10^5 bacteria of the strains to be tested. At 15 and 21 days after inoculation, six plantlets were removed from Gibson tubes and placed together in airtight tubes. Then, 1 mL of acetylene was added to the tubes and the plantlets were incubated at 28°C during 4 h. After this time, a volume of 0.4 mL of the gas present in the tube was directly injected on a gas chromatograph (Agilent GC7820) and the area of the ethylene peak was measured. Empty tubes were used as negative control during the assays as well as tubes containing plantlets without acetylene addition. Plants inoculated with the wild-type strain of *C. taiwanensis* LMG19424 was used as positive control and nitrogen-fixing reference strain. All data were analyzed with the GC OpenLAB application.

Infectivity assays

Infectivity assays were performed as described by (Doin de Moura et al., 2022). Briefly, *Mimosa pudica* plantlets cultivated in tubes containing 20 mL of solid Fahreaus medium and 40 mL of liquid Jensen medium diluted 1:4 with sterile water were singly inoculated with each bacterial strain to be tested. Plants inoculated with each strains were incubated in a culture chamber at 28°C under a photoperiod day/night of 16h/8 h. Fifteen days after inoculation, nodules from each plant (6 plants/strain/assay) were harvested, counted and crushed in sterile water, diluted and spread on selective solid medium using an easySpiral automatic plater (Interscience). Colonies from each strain were counted, allowing the measurement of the number of bacteria per nodule. Pairwise comparison of infectivity values were performed using Wilcoxon-rank sum test ($p<0.05$).

Sup Table 4.1. Strains used in this study

Resource type	Designation	Reference/source	Identifier	Additional information/Common name
Cupriavidus taiwanensis	Wild-type strain isolated from *Mimosa pudica* in Taiwan	Chen et al., 2001	LMG19424	
Cupriavidus taiwanensis	LMG19424 derivative resistant to StrR	M. Hynes	CBM832	
Ralstonia solanacearum	Non-nodulating chimeric ancestor, TriR	Marchetti et al., 2010	CBM124	GMI1000 pRalta::TriR
Ralstonia solanacearum	Nodulating ancestor mutated in *hrpG* , TriR, GenR	Marchetti et al., 2010	CBM212	CBM212
Ralstonia solanacearum	Nodulating ancestor mutated in *hrpG* , TriR, GenR	Marchetti et al., 2010	CBM349	CBM349
Ralstonia solanacearum	Nodulating ancestor mutated in *hrcV* , TriR, GenR	Marchetti et al., 2010	CBM356	CBM356
Ralstonia solanacearum	Evolved clone from line B, cycle 16, TriR, GenR	Marchetti et al., 2010	RCM1492	Clone B16
Ralstonia solanacearum	Evolved clone from line B, cycle 35, TriR, GenR	Doin de Moura et al., 2022	RCM2458	Clone B35
Ralstonia solanacearum	Evolved clone from line B, cycle 38, TriR, GenR	This work	RCM2870	Clone B38
Ralstonia solanacearum	Evolved clone from line B, cycle 39, TriR, GenR	This work	RCM2890	Clone B39
Ralstonia solanacearum	Evolved clone from line B, cycle 40, TriR, GenR	This work	RCM2909	Clone B40
Ralstonia solanacearum	Evolved clone from line B, cycle 41, TriR, GenR	This work	RCM2927	Clone B41
Ralstonia solanacearum	Evolved clone from line B, cycle 42, TriR, GenR	This work	RCM2954	Clone B42
Ralstonia solanacearum	Evolved clone from line B, cycle 44, TriR, GenR	This work	RCM3124	Clone B44
Ralstonia solanacearum	Evolved clone from line B, cycle 45, TriR, GenR	This work	RCM3145	Clone B45
Ralstonia solanacearum	Evolved clone from line B, cycle 46, TriR, GenR	This work	RCM3175	Clone B46
Ralstonia solanacearum	Evolved clone from line B, cycle 50, TriR, GenR	This work	RCM3300	Clone B50
Ralstonia solanacearum	Evolved clone from line B, cycle 55, TriR, GenR	This work	RCM3457	Clone B55
Ralstonia solanacearum	Evolved clone from line B, cycle 60, TriR, GenR	This work	RCM3634	Clone B60
Ralstonia solanacearum	Evolved clone from line F, cycle 16, TriR, GenR	Marchetti et al., 2010	RCM2182	Clone F16
Ralstonia solanacearum	Evolved clone from line F, cycle 25, TriR, GenR	Doin de Moura et al., 2022	RCM2206	Clone F25
Ralstonia solanacearum	Evolved clone from line F, cycle 30, TriR, GenR	Doin de Moura et al., 2022	RCM2339	Clone F30
Ralstonia solanacearum	Evolved clone from line F, cycle 36, TriR, GenR	This work	RCM2872	Clone F36
Ralstonia solanacearum	Evolved clone from line F, cycle 38, TriR, GenR	This work	RCM2911	Clone F38
Ralstonia solanacearum	Evolved clone from line F, cycle 50, TriR, GenR	This work	RCM3301	Clone F50
Ralstonia solanacearum	Evolved clone from line F, cycle 60, TriR, GenR	This work	RCM3635	Clone F60
Ralstonia solanacearum	Evolved clone from line G, cycle 25, TriR, GenR	Doin de Moura et al., 2022	RCM2162	Clone G25
Ralstonia solanacearum	Evolved clone from line G, cycle 26, TriR, GenR	Doin de Moura et al., 2022	RCM2207	Clone G26
Ralstonia solanacearum	Evolved clone from line G, cycle 27, TriR, GenR	Doin de Moura et al., 2022	RCM2234	Clone G27
Ralstonia solanacearum	Evolved clone from line G, cycle 28, TriR, GenR	Doin de Moura et al., 2022	RCM2257	Clone G28
Ralstonia solanacearum	Evolved clone from line G, cycle 30, TriR, GenR	Doin de Moura et al., 2022	RCM2307	Clone G30
Ralstonia solanacearum	Evolved clone from line G, cycle 35, TriR, GenR	Doin de Moura et al., 2022	RCM2460	Clone G35
Ralstonia solanacearum	Evolved clone from line G, cycle 38, TriR, GenR	This work	RCM2871	Clone G38
Ralstonia solanacearum	Evolved clone from line G, cycle 39, TriR, GenR	This work	RCM2891	Clone G39
Ralstonia solanacearum	Evolved clone from line G, cycle 40, TriR, GenR	This work	RCM2910	Clone G40
Ralstonia solanacearum	Evolved clone from line G, cycle 50, TriR, GenR	This work	RCM3302	Clone G50
Ralstonia solanacearum	Evolved clone from line G, cycle 55, TriR, GenR	This work	RCM3459	Clone G55
Ralstonia solanacearum	Evolved clone from line G, cycle 60, TriR, GenR	This work	RCM3636	Clone G60
Ralstonia solanacearum	Evolved clone from line K, cycle 16, TriR, GenR	Marchetti et al., 2010	RCM2184	Clone K16
Ralstonia solanacearum	Evolved clone from line K, cycle 25, TriR, GenR	Doin de Moura et al., 2022	RCM2208	Clone K25
Ralstonia solanacearum	Evolved clone from line K, cycle 36, TriR, GenR	This work	RCM2873	Clone K36
Ralstonia solanacearum	Evolved clone from line K, cycle 38, TriR, GenR	This work	RCM2912	Clone K38
Ralstonia solanacearum	Evolved clone from line K, cycle 50, TriR, GenR	This work	RCM3303	Clone K50
Ralstonia solanacearum	Evolved clone from line K, cycle 60, TriR, GenR	This work	RCM3637	Clone K60
Ralstonia solanacearum	Evolved clone from line M, cycle 36, TriR, GenR	This work	RCM2874	Clone M36
Ralstonia solanacearum	Evolved clone from line M, cycle 38, TriR, GenR	This work	RCM2913	Clone M38
Ralstonia solanacearum	Evolved clone from line M, cycle 40, TriR, GenR	This work	RCM2958	Clone M40
Ralstonia solanacearum	Evolved clone from line M, cycle 46, TriR, GenR	This work	RCM3179	Clone M46
Ralstonia solanacearum	Evolved clone from line M, cycle 49, TriR, GenR	This work	RCM3270	Clone M49
Ralstonia solanacearum	Evolved clone from line M, cycle 50, TriR, GenR	This work	RCM3304	Clone M50
Ralstonia solanacearum	Evolved clone from line M, cycle 55, TriR, GenR	This work	RCM3461	Clone M55
Ralstonia solanacearum	Evolved clone from line M, cycle 60, TriR, GenR	This work	RCM3638	Clone M60
Ralstonia solanacearum	Evolved clone from line X, cycle 1, TriR, GenR	This work	RCM3181	Clone X01
Ralstonia solanacearum	Evolved clone from line X, cycle 2, TriR, GenR	This work	RCM3243	Clone X02
Ralstonia solanacearum	Evolved clone from line X, cycle 3, TriR, GenR	This work	RCM3272	Clone X03
Ralstonia solanacearum	Evolved clone from line X, cycle 6, TriR, GenR	This work	RCM3373	Clone X06
Ralstonia solanacearum	Evolved clone from line X, cycle 10, TriR, GenR	This work	RCM3498	Clone X10
Ralstonia solanacearum	Evolved clone from line X, cycle 15, TriR, GenR	This work	RCM3647	Clone X15
Ralstonia solanacearum	Evolved population from line B, cycle 37, TriR, GenR	This work	RCM2852	Evolved population B37
Ralstonia solanacearum	Evolved population from line B, cycle 39, TriR, GenR	This work	RCM2878	Evolved population B39
Ralstonia solanacearum	Evolved population from line B, cycle 42, TriR, GenR	This work	RCM2942	Evolved population B42
Ralstonia solanacearum	Evolved population from line B, cycle 43, TriR, GenR	This work	RCM3089	Evolved population B43
Ralstonia solanacearum	Evolved population from line B, cycle 45, TriR, GenR	This work	RCM3133	Evolved population B45
Ralstonia solanacearum	Evolved population from line B, cycle 47, TriR, GenR	This work	RCM3183	Evolved population B47
Ralstonia solanacearum	Evolved population from line M, cycle 37, TriR, GenR	This work	RCM2882	Evolved population M37
Ralstonia solanacearum	Evolved population from line M, cycle 39, TriR, GenR	This work	RCM2919	Evolved population M39
Ralstonia solanacearum	Evolved population from line M, cycle 41, TriR, GenR	This work	RCM3019	Evolved population M41
Ralstonia solanacearum	Evolved population from line M, cycle 43, TriR, GenR	This work	RCM3093	Evolved population M43
Ralstonia solanacearum	Evolved population from line M, cycle 45, TriR, GenR	This work	RCM3137	Evolved population M45
Ralstonia solanacearum	Evolved population from line M, cycle 47, TriR, GenR	This work	RCM3187	Evolved population M47
Ralstonia solanacearum	Evolved population from line X, cycle 1, TriR, GenR	This work	RCM3165	Evolved population X01
Ralstonia solanacearum	Evolved population from line X, cycle 3, TriR, GenR	This work	RCM3256	Evolved population X03
Ralstonia solanacearum	Evolved population from line X, cycle 5, TriR, GenR	This work	RCM3320	Evolved population X05
Ralstonia solanacearum	Chimeric ancestor carrying the mutation *hrpG* -Q81* from the nodulating ancestor CBM212, TriR	Marchetti et al., 2010	CBM1627	CBM124 *hrpG* -Q81*
Ralstonia solanacearum	Reconstructed strain carrying the mutations *hrpG* -Q81* and *efpR* -E66K ,TriR	This work	RCM1865	CBM124 hrpG-Q81* *efpR* -E66K
Ralstonia solanacearum	Reconstructed strain carrying the mutations *hrpG* -Q81* and RSc2277-V321G ,TriR	This work	RCM3467	CBM124 hrpG-Q81* RSc2277-V321G
Ralstonia solanacearum	Reconstructed strain carrying the mutations *hrpG* -Q81*, *efpR* -E66K and RSc2277-V321G ,TriR	This work	RCM3475	CBM124 hrpG-Q81* *efpR* -E66K RSc2277-V321G
Ralstonia solanacearum	Evolved clone from line G, cycle 1, TriR, GenR	Marchetti et al., 2010	CBM420	Clone G1
Ralstonia solanacearum	Evolved clone from line G, cycle 4, TriR, GenR	Marchetti et al., 2010	CBM557	Clone G4
Ralstonia solanacearum	Evolved clone from line G, cycle 11, TriR, GenR	Marchetti et al., 2010	RCM1281	Clone G11
Ralstonia solanacearum	Evolved clone from line G, cycle 25, TriR, GenR	Marchetti et al., 2010	RCM2162	Clone G25
Ralstonia solanacearum	Evolved clone G1 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM1818	Clone G1 IGglmS:P*psbA* -GFP

Ralstonia solanacearum	Evolved clone G4 carrying a P*psbA* -GFP fusion downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM1958	Clone G4 IGglmS:P*psbA* -GFP
Ralstonia solanacearum	Evolved clone G1 *efpR* -E66K carrying a KanR cassette downstream glmS, TriR, KanR	Capela et al., 2017	RCM1867	G1 *efpR* -E66K IGglmS:KanR
Ralstonia solanacearum	Evolved clone G4 *ppiB* -Q112* carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3084	G4 *ppiB* -Q112* IGglmS:P*psbA* -mcherry
Ralstonia solanacearum	Evolved clone G4 RSc2277-V321G carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3083	G4 RSc2277-V321G IGglmS:P*psbA* -mcherry
Ralstonia solanacearum	Evolved clone G11 RSc0194-S102F carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3111	G11 RSc0194-S102F IGglmS:P*psbA* -mCherry
Ralstonia solanacearum	Evolved clone G25 RSp0759-P92S carrying a P*psbA* -mCherry fusion and a KanR cassette downstream glmS, TriR, KanR	Doin de Moura et al., 2022	RCM3087	G25 RSp0759-P92S IGglmS:P*psbA* -mCherry

TriR, trimethoprim resistance. KanR, kanamycin resistance. StrR, streptomycin resistance. IGglmS: insertion in the intergenic region downstream the *glmS* gene.

CHAPTER 5

General conclusion and perspectives

From an evolutionary point of view, rhizobia constitute a group of bacteria particularly interesting to study. Indeed, these bacteria able to establish a complex interaction with legume plants are phylogenetically diverse, scattered in 18 or 19 different genera among two or three classes of proteobacteria and intermixed with non-rhizobium species (Masson-Boivin et al., 2009; Tang and Capela, 2020). Field experiments and phylogenetic studies have established that horizontal transfers of key symbiotic genes (mainly the *nod*, *nif* and *fix* genes) have played a foundational role in the emergence and diversification of rhizobia, at least for the Nod Factor-dependent ones (Andrews et al., 2018; Barcellos et al., 2007; Nandasena et al., 2007; Novikova and Safronova, 1992; Pérez Carrascal et al., 2016; Rogel et al., 2001; Sullivan et al., 1995). However, although horizontal transfers occur frequently in nature, transfers successfully converting a recipient bacterium into an efficient legume symbiont were probably rare and likely dependent on particular circumstances such as the introduction of a legume species in a new location, the presence of locally adapted soil bacteria and the absence of competitive associated symbionts. Moreover, taking into account the diversity of microorganisms in the soil, the relatively small number of genera containing rhizobia, raises the question of a required genetic predisposition to be converted in legume symbiont. Adaptation to the rhizosphere and acidic conditions, common features of rhizobia, have been suggested by the studies of (Garrido-Oter et al., 2018) and (Fagorzi et al., 2020) to be part of these predisposing functions. In addition to the putative required predisposing functions, a body of data suggest that horizontal transfers were followed by post-transfer adaptation steps. Indeed, even if laboratory experiments have shown that in some cases the transfer of symbiotic plasmids can directly convert a non-nodulating recipient bacterium into a N_2-fixing legume symbiont (Rogel et al., 2001; Suominen et al., 2001), in most cases these transfers are unsuccessful or lead to suboptimal rhizobia (Abe et al., 1998; Hirsch et al., 1984b; Nakatsukasa et al., 2008; Nandasena et al., 2007). For example, one crucial step in the activation of the symbiotic potential may be the recruitment of essential nodulation and nitrogen fixation genes from endogenous housekeeping genes (when they are absent on symbiotic plasmids or islands) either directly (*glmS* can replace *nodM*, *iscA* and *iscS* can replace *nifU* and *nifS*, respectively) or following a duplication and integration into the symbiotic region (*glmS*/*nodM* or *cysDNC*/*nodPQ*) (Galibert et al., 2001; Schwedock & Long, 1992). The post-transfer adaptation hypothesis is also supported by the fact that the genes involved in symbiosis are very diverse from a rhizobium to another, many of them being lineage-specific (Amadou et

al., 2008; Tian, 2012). In some cases, rhizobia have used very specific metabolic properties, such as methylotrophy in *Methylobacterium nodulans* or photosynthesis in *Bradyrhizobium* sp. ORS278, to perform symbiosis (Giraud et al., 2000; Jourand et al., 2005). Other possible post-transfer adaptive steps may include allelic variation, inactivation or deletion of incompatible functions, or additional horizontal gene acquisition (Masson-Boivin et al., 2009). Finally, after the activation of their symbiotic capacities, rhizobia have certainly optimized their mutualistic relationship via a permanent remodeling of their genome over the years of coevolution with their hosts.

Although strongly suggested by these data, the post-transfer adaptation stages have not been fully demonstrated. This is what we wanted to validate through the evolution experiment carried out in the team. In this experiment, the transfer of the *nod, nif and fix* genes from *Cupriavidus taiwanensis,* the symbiont of *Mimosa pudica*, to the plant pathogen *Ralstonia solanacearum* did not immediately generate a symbiont but the early steps of the symbiosis, *i.e.* nodulation and intracellular infection of nodules, were rapidly activated following the inactivation of incompatible functions, mainly related to virulence (Marchetti et al., 2010). Previous works in the team showed that these steps were then strongly improved through the rewiring of the virulence regulatory network of the recipient bacterium, thus validating, in this case, the need of additional adaptation steps after the transfer of symbiotic genes (Capela et al., 2017; Guan et al., 2013; Tang et al., 2020). Moreover, these works have also shown that multiple genetic routes can lead to similar phenotypes, confirming the high plasticity of bacterial genomes to adapt to new environments. In this thesis, we went further in the understanding of the adaptation process of bacteria in this experiment. By using a new approach based on the sequencing of entire evolved populations, we have been able to analyze the dynamics of bacterial adaptation, observe the selective forces acting in this system and identify the biological functions potentially required or deleterious for the symbiotic process. Here, I discuss the main results of this thesis, as well as the remaining questions to be answered in future works.

1. Analysis of the evolved population by a metagenomic approach: a powerful tool to analyze the adaptation process

Experimental evolution is a technique that allows to study adaptive processes in the laboratory, and that is especially well suited for microorganisms with large populations and

fast generation times. The analysis of experimentally evolved populations can help uncover the genetic bases of phenotypes of interest (the nitrogen fixing symbiosis in our case) and it can also shed light on the dynamics of adaptive processes.

The advent of new sequencing technologies has greatly helped with the identification of the genetic bases of adaptation, since it is now possible to sequence the entire genome of tens to hundreds of evolved bacterial clones in order to discover which mutations are responsible for the phenotypic changes. This approach has been used previously in the group but has shown some limitations. Indeed, due to a phenomenon of transient hypermutagenesis occurring in the rhizosphere of the plants during the evolution experiment (Remigi et al., 2014), the high number of mutations acquired in each evolution cycle makes the identification of adaptive mutations difficult. In this case, the use of whole-population metagenomic sequencing represents an interesting tool to select mutations based on their trajectories (their changes in frequency over time). This way, we can hypothesize that a mutation increasing in frequency in the evolved populations confers an adaptive advantage to the clones carrying it. Using this rationale to select potentially-adaptive mutations, I was able to identify 25 adaptive mutations in 2 independent lineages. Moreover, the metagenomics study of the evolved populations in our system gave us a comprehensive view of the molecular events that occurred during the experiment and the dynamics of the adaptation. Firstly, by following the trajectories of nearly all mutations from these populations, we observed events of rapid selective sweeps and clonal interference, similar to what was described in more simple systems (Cvijović et al., 2018; Good et al., 2017). Less common was the presence of mutational cohorts composed by up to 30 mutations following the same frequencies over time (Lang et al., 2013) probably arising due to hitchhiking (i.e. when a neutral or a slightly deleterious mutation increases in frequency as a consequence of being in the same group of an adaptive mutation). Interestingly, we have found that some large cohorts harbor several adaptive mutations in our system. While our naïve expectation was that the rise in frequency of each cohort would be driven by one adaptive mutation, this phenomenon is possibly the sign of intense clonal interference (due to high mutation rate and the diversity of adaptive mutations) occurring in our system. Indeed, it is possible that a given clone remains at low frequency (i.e. beyond our detection limit of 5%) in our populations while gradually accumulating adaptive mutations, until its fitness significantly exceeds that of all other clones present in the population, at which point it will increase in frequency. In this case, the deep sequencing of tagged populations could allow us

to observe details about the dynamics of these mutations, as performed in yeast by (Nguyen Ba et al., 2019). Additionally, the existence of epistatic interactions between the genes affected by the adaptive mutations from a same cohort remains to be investigated.

The hypermutagenesis occurring in the experiment allows a rapid adaptation of bacteria to symbiosis, despite the very strong bottleneck acting at the root entry point. Interestingly, no sign of a reduced hypermutagenesis was observed in the experiment, even after 60 evolution cycles. Because the majority of mutations are expected to be deleterious (Bataillon and Bailey, 2014), a decrease in mutation rate would likely be beneficial when bacteria approach their fitness optimum. Such a decrease in mutation rates in hypermutable populations was observed in several evolution experiments (Chen and Zhang, 2014; Lynch et al., 2016; McDonald et al., 2012; Raynes & Sniegowski, 2014; Wielgoss et al., 2013). However, the phenomenon of transient hypermutagenesis is particular in the sense that mutation rate is dependent on the level of stress experienced by the bacteria: if bacteria adapt to the stressful conditions and/or spend less time in the rhizosphere because they form nodules more rapidly, then they should accumulate fewer mutations. The other peculiarity of our system is the strong selective bottleneck at the nodulation step, which may rapidly eliminate clones carrying deleterious mutations. Together, these two effects probably limit the efficiency of selection favoring lower mutation rates. To investigate if hypermutagenesis is counter-selected in our experimental conditions, it would be interesting to analyse the fitness effects (both positive and negative) of hypermutagenesis at different times of the evolution experiment, by competing evolved clones with their isogenic non-hypermutable mutants over multiple nodulation cycles.

2. Nodulation competitiveness: a symbiotic trait driving the emergence of new symbionts?

In this work, we observed the fast and parallel improvement of nodulation competitiveness in all evolved lineages. Accordingly, all the adaptive mutations that we identified improved this trait, while only a subset of these mutations also improved infectivity. These results suggest that selection for nodulation competitiveness is the major force driving adaptation to symbiosis. Evolutionary simulations support this conclusion, and the selective nodulation bottleneck appeared as an important factor responsible for this effect. However, these simulations were based on the hypothesis that the distribution of fitness effects is similar for the 2 traits analyzed here (nodulation competitiveness and infectivity). Testing this hypothesis

would require to measure the symbiotic phenotypes of a large number of random mutants. Such studies were performed for simple phenotypes (such as antibiotic resistance (Kassen and Bataillon, 2006; Lundin et al., 2018), but would be extremely labor intensive (and nearly impossible) for symbiotic phenotypes. Therefore, while we have not found any mutation that improve only infection, we cannot distinguish between 2 possibilities: either these mutations occur in the evolution cycles but are not selected, or these mutations simply do not exist or are extremely rare compared to mutations affecting nodulation. However, it is possible that such mutations will be identified as we continue the evolution experiment, possibly using a different selective regime to promote the selection of nitrogen fixation (see paragraph 4 below).

We also evidenced experimentally a genetic coupling between nodulation and intracellular infection of nodule cells. This coupling was suggested by (Guan et al., 2013), but demonstrated in this work on several different mutants. The coupling between nodulation and intracellular infection observed in this evolution experiment raises interesting questions about their mechanistic bases. We could hypothesize that the same biological functions are involved in both phenotypes. Contrariwise, it is possible that selected mutations (which very often presented pleiotropic effects) modify several functions involved in one or other phenotype individually. The genetic dissection of biological functions involved in the different steps of the interaction (see paragraph 3 below) should help resolve this question. In addition, we can wonder whether other symbiotic traits, namely the ones related to late symbiotic steps, were also linked to the initial selection for nodulation at the root entry point. An adaptive mutation identified in this work, RSc2277-V321G supports this hypothesis. This mutation improves strongly the nodulation and intracellular infection, but also allows a first level of nitrogenase activity. Besides that, in several lineages, mutations in this gene were present in clones showing shifts in nitrogenase activity. At this point, we do not know what portion of the possible adaptive mutations affect one single or several phenotypic trait(s) or even whether there are examples of trade-offs between the phenotypes. It was well described that nodulation is affected by several functions. However, not much is known about the functions necessary to improve intracellular infection of nodule cells. Further investigating these points will be very important to fully understand the evolution and functioning of nitrogen-fixing symbiosis. In this respect, it would be interesting to have a full phenotypic characterization of the myriad of rhizobial mutants that were described in the literature.

3. Adaptive evolution is mainly driven by pleiotropic mutations affecting convergent molecular functions

A major goal in the team is to understand the molecular functions involved in symbiotic adaptation. In this work, we have identified a pool of new mutations improving different symbiotic traits (i.e. nodulation competitiveness, within-host proliferation and nitrogenase activity). However, most of these mutations affect uncharacterized or poorly characterized genes and rarely point to one specific biological function, except in some cases, such as the T3SS (*hrcV*) or the T6SS (*tssJ*, RSp0759, RSp0769), although the latter remains to be fully characterized. Moreover, as often in experimental evolution, the characterization of the most adaptive mutations indicated that they target genes with pleiotropic effects. Interestingly, in the chapter 3 of this thesis, we showed that several of these strongly adaptive, apparently unrelated, mutations do affect some common biological functions, including motility, EPS biosynthesis, metabolic properties and stress tolerance. Convergence in terms of genes, pathways and functions is also commonly observed in experimental evolution, either on independent different replicate lines or in different sub-populations co-existing in a given line, and is generally a strong indication of involvement in adaptation (Tenaillon et al., 2012). The probability of convergent changes seems to depend on the level of biological organization, meaning that the chances to converge increase as we move from single bases, to genes, to pathways and to biological functions. In this experiment, we observed very rare cases of convergence at the level of base pair substitutions (except the mutation *efpR*-E66K observed in three independent lineages (Capela et al., 2017) while several examples of convergences at the levels of genes (*hrpG*, *efpR,* RSc2277) and pathways (*hrpG*, *efpR*, *phc*) were described in previous works or in this thesis (Capela et al., 2017; Guan et al., 2013; Marchetti et al., 2010; Tang et al., 2020). The functions convergently affected by the newly identified mutations need now to be further investigated to assess their involvement in symbiotic adaptation. Among these functions, the better tolerance to acidic pH, some enhanced catabolic properties, in particular for amino acids and dicarboxylic acids, and the increase in motility, are good candidates to play a role in symbiosis. Indeed, these functions are related to the environmental conditions likely encountered by bacteria during the symbiotic process. For example, acidic pH may be present in the rhizospheric compartment as well as in symbiosomes where plant H^+-ATPase act. Bacterial motility may enable a better access to the host rhizosphere. Catabolic activities may help bacteria to assimilate more efficiently nutrients

available in the rhizosphere and in nodule cells. The reduction in EPS production could have been another good candidate function involved in symbiotic adaptation in particular to avoid plant recognition during infection. However, so far, (Tang et al., 2020) showed that an *eps* defective mutant is not adaptive for symbiosis. In future work, it will be important to modify these functions individually and/or in combination with the other functions to understand how bacteria adapt to symbiosis and whether several functions need to be modified at the same time to allow the achievement of certain stages of symbiosis. In addition, it will be interesting to understand how these different adaptive mutations change these biological functions and determine their role or position in the gene regulatory network of *Ralstonia*. Finally, our results present experimental evolution as an efficient way to discover the function of new genes, as well as the connections between known pathways.

4. The evolution of mutualism was not elucidated

In this evolution experiment, mutualism was not obtained showing that the emergence of new legume symbionts is challenging even when the transfer of the symbiotic genes occurred between closely related strains belonging to neighboring genera. Two possible issues may explained this result, the genetic background of the recipient strain and the selective regime used in this experiment.

The key step that seems to block the evolutionary process towards a functional symbiosis is the persistence of bacteria once released in symbiosomes into the plant cells. Symbiosomes are indeed special compartments where the oxygen content is very low and strictly controlled by the plant *via* the production of leghemoglobins and where available nutrients are only those delivered by the plants. In addition, bacterial divisions and basic metabolism are blocked in these compartments again probably by plant control mechanisms. These mechanisms were relatively well described in IRLC legumes and Dalbergioid such as *Medicago* and *Aeschynomene spp.,* with the expression of Nodule specific Cysteine-Rich (NCR) peptides*,* but are completely unknown in most legumes. In terms of genetics, we may wonder whether there are some genes or functions essential for late symbiotic stages in *C. taiwanensis* that are missing in *R. solanacearum.* Since *Mimosa pudica* is not a natural host plant of *R. solanacearum*, it is possible that specific (pre-adaptive?) bacterial functions necessary to interact and colonize this plant species are absent in *R. solanacearum*. A fine comparison of the genomes and the metabolic capacities of *R. solanacearum* and *C. taiwanensis* coupled to

metabolomics of nodules could help identifying some missing functions. Among the crucial late symbiotic functions, genes involved in bacterial respiration under stringent microoxic conditions are critical. Although the genome of *R. solanacearum* encodes one *cbb*3 cytochrome oxidase, it may not be fully adapted to allow bacterial respiration in a nodule environment. Also an ammonium transport system is present in the genome of *R. solanacearum* but the regulation to switch from ammonium assimilation to ammonium export may not be in place in this bacterium. Finally, although no NCR peptides were identified in the *Mimosa* genome and the bacterial counterpart *bacA* is absent in *C. taiwanensis*, the existence of a mechanism controlling bacteroid differentiation in *Mimosa* cannot be excluded. Interestingly, the expression of many small proline-rich peptides was recently detected in *Mimosa* nodules (Libourel et al., 2022). However, their role in symbiosis remains to be investigated. One way to identify genes that might be missing in *Ralstonia* to achieve mutualism would be to introduce specific candidate genes or random genomic DNA fragments of *C. taiwanensis* into a *Ralstonia* evolved clone already well adapted to early symbiotic stages and then screen the complemented strains for their ability to promote plant growth. The fact that evolved *R. solanacearum* failed to form N_2-fixing nodules on *Mimosa* thus give us the opportunity to discover the symbiotic mechanisms at play during the late stages of symbiosis in Mimosoideae. These mechanisms are so far unknown and probably different from other model legume species.

Besides genes that may be missing, the selective regime used in the evolution experiment may have been another obstacle to the emergence of N_2-fixing clones. Indeed, the selection for nitrogen fixation was shown to be stronger in longer nodulation cycles (Daubech et al., 2017). It is possible that the 21-day cycles applied in this evolution experiment were not optimal to favor the increase in frequency of nitrogen-fixing clones in bacterial populations. In line with this remark, in nature, rhizobia are exposed to nodulation cycles of several months. Increasing the length of the evolution cycles could have been a good way to select nitrogen fixers once nodulation and intracellular infection were optimized. Other ecological parameters in our experimental set up could have been modified to increase the chances of success in obtaining N_2-fixers. In particular, increasing the number of plants inoculated in each evolution cycle, and consequently increasing the number of nodules, could have helped to trap rare N_2-fixing variants.

Another strategy would have been to make an artificial selection or a guided selection of nodules either based on their color and size or based on their effect on plant growth and color. This strategy was successfully used in other experiments to select insect mutualists or microbial communities with beneficial effects on plant or animal health (Koga et al., 2022; Mueller and Sachs, 2015). Indeed, natural selection of bacteria maximizing their fitness, as in our experiment, may not be the best way to obtain mutualists. Bacteria can find strategies to increase their fitness in nodules without providing any benefits to the host plant or by using the nitrogen fixed for their own growth. Although we did not see it in the experiments from the chapter 3 of this thesis, it is not impossible that in this experiment we select for *Mimosa* parasites and thus we might be in an evolutionary impasse to achieve mutualism.

5. Is it possible to convert a pathogen into symbiont?

Pathogenic and mutualistic microbes have traditionally been studied completely independently. However, both phylogenetic inferences and direct experimental evolution have shown that transitions between these two lifestyles are not uncommon. It is now clear that mutualism and pathogenicity are evolutionary and functionally related, and represent two extremes of a continuum rather than separate phenomena (Drew et al., 2021; Flórez et al., 2017; Sachs et al., 2014; Tso et al., 2018). An important question in the field is to identify what biological functions are common or specific to these lifestyles (beyond the well-known virulence or mutualistic traits).

Our experimental system is a good model to address this question since we aim to convert a plant pathogenic bacterium into a legume mutualist. Although the transition is still incomplete since the evolved strains behave as commensals, we can begin to analyze the genetic bases of adaptation to the first steps of this symbiotic interaction. Among the biological functions required for *R. solanacearum*'s pathogenicity, some may be required to sustain life within the host plant or survival to stress conditions imposed by the host, such as high osmolarity and acidic pH, while others could be detrimental for symbiosis, for example if they elicit plant immunity. Therefore, identifying which functions are conserved, improved or lost during adaptation to symbiosis should help uncover the functional links between these 2 lifestyles. In our experiment, almost all mutations that are strongly adaptive for symbiosis also decreased pathogenicity on *Arabidopsis thaliana*. However, because all of these adaptive mutations (except RSp0595-P132P) are pleiotropic, it is unclear if pathogenicity loss is directly

linked to improved symbiosis, or if it is a side effect of relaxed selection on pathogenicity determinants. Here again, the genetic dissection of the different biological functions affected by these mutations should help answer this question.

Although transitions from pathogenicity to mutualism occurred in many bacterial clades, no example support such transitions in rhizobia (Sachs et al., 2014). We can wonder whether the lack of transitions from pathogens to rhizobia is linked to strong immune responses in the nodules. Indeed, in addition to classical PTI and ETI defense responses found in all plant organs, nodules establish specific intracellular defense responses to control the proliferation of symbiotic bacteria within plant cells (Gourion et al., 2015). The first adaptive mutations (in *hrcV* and *hrpG*) allowing nodulation disrupted the activity of the T3SS, suggesting that the recognition of one or several T3Es blocked nodulation, likely due to ETI. In addition, symbiotic adaptation until cycle 16 was accompanied by the disappearance of necrotic-like areas within nodules and of the ability of *Ralstonia* strains to elicit immune responses on a compatible host plant (Marchetti et al., 2014). These results indicate that PTI and ETI represented strong barriers to the establishment of symbiosis with a pathogenic bacterium, but that these barriers were rapidly overcome during experimental evolution. The third layer of defense responses, acting on intracellular bacteria, has been characterized to some extent in model legumes, in particular in *Medicago truncatula* with the discovery of NCR peptides and other defense-related genes such as *DNF2*, *NAD1, SymCRK* or *RSD* (Benezech et al., 2020). These defense mechanisms seem to show significant variability between the different plant lineages, and the mechanisms allowing the control of intracellular bacteria in *Mimosa pudica* are still unknown (beyond the recent discovery that small proline-rich peptides might be involved in this process (Libourel et al., 2022). Since plants from this genus are able to develop symbiotic interactions with different genera of rhizobia, namely *Rhizobium*, *Paraburkholderia* and *Cupriavidus* (Klonowska et al., 2012), it is possible that the genes required to survive within these plants may be shared by multiple bacterial genera. Unfortunately, *Ralstonia* does not seem to be one of those.

In summary, we propose that the transition from pathogens to rhizobia may require 2 steps (following the acquisition of the symbiotic genes). The first step is the loss of major pathogenicity determinants and immune-activating factors that can trigger PTI or ETI. This step leads to a non-functional symbiosis (commensalism) and seems relatively easy, since it occurred in multiple independent lineages in our experiments. This may also explain why a

number of non-rhizobial bacteria are able to colonize nodules as commensals (Martínez-Hidalgo and Hirsch, 2017; Mayhood and Mirza, 2021). The second step is the acquisition/actualization of mutualism, that encompasses (i) the evasion of and/or tolerance to nodule-specific immune responses, (ii) the adaptation to the intracellular niche (including low pH, low oxygen concentration and nutrient availability) and (iii) the reduction of nitrogen and export of ammonia to the plant. This step is probably more stringent, since legume plants sometimes produce a vast array of antimicrobial peptides, and intracellular proliferation is unusual for plant-associated bacteria. Therefore, this step may rely on pre-adaptation of bacterial genomes (Fagorzi et al., 2020), and thus contribute to explain why rhizobia are restricted to few genera (despite massive horizontal gene transfers).

6. Rhizobia have evolved in complex eco-evolutionary systems:

In this experiment, the *Ralstonia* clones have evolved under a simple system. In terms of biological diversity, the bacteria interact with only one bacterial and one plant species in each evolution cycle. These conditions do not reconstitute completely what rhizobia face in natural bulk soils, in the rhizosphere of plants and in nodule of natural partners. Additionally, in terms of environmental conditions, the evolution experiment design exposes bacteria to a very simplified system, where plants are grown in a semi-sterile synthetic medium with controlled temperature and light, and nodules are harvest before natural senescence. In contrast, natural interactions between rhizobia and leguminous plants expose them to complex biotic environments at each step of the symbiosis, involving interactions with many species composing the soil, rhizospheric, and nodule microbiomes. For example, in the soil, rhizobia interact constantly with other bacterial and non-bacterial (phages, fungi, ciliates, nematodes, etc) species. In this step, the survival of rhizobia requires the ability to adapt to specific environmental conditions, as well as to bypass negative ecological interactions, such as competition, antibiosis and predation. This means that the free-living step of the life cycle of rhizobia may also select bacteria for a better competition in the soil, such as antimicrobial activities in order to avoid competitors. Besides that, in some cases, the evolution of the rhizobia species occurred during the interaction with several plant species or plants showing intra-species diversity. In these cases, we can wonder whether several events of acquisition of different symbiotic genes were necessary to the adaptation to the mentioned diverse range of hosts. Possibly, the adaptation to different hosts also requires the activation or the

repression of various functions, meaning that the adaptation to those plants can involve several post-transfer adaptation steps. Taking all this into account, it is important to remind that, although this EE brings important insights on bacterial adaptation to symbiosis with legumes, it remains a simplified system and in nature other complex selective forces may be acting. In future experimental evolution works, the use of more than one bacterial species, natural soil or multiple hosts could allow the team to further understand the evolution and emergence of new legume symbionts in the wild.

References

Abe, M., Kawamura, R., Higashi, S., Mori, S., Shibata, M., & Uchiumi, T. (1998). Transfer of the symbiotic plasmid from *Rhizobium leguminosarum* biovar *trifolii* to *Agrobacterium tumefaciens*. **Journal of General and Applied Microbiology**, 44(1), 65-74.

Amadou, C., Pascal, G., Mangenot, S., Glew, M., Bontemps, C., Capela, D., Masson-Boivin, C. (2008). Genome sequence of the beta-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. **Genome Research**, 18(9), 1472-1483. https://doi.org/10.1101/gr.076448.108

Amarger, N. (1981). Competition for nodule formation between effective and ineffective strains of R*hizobium meliloti*. **Soil Biology and Biochemistry**, v.13.

Andrews, M., De Meyer, S., James, E. K., Stępkowski, T., Hodge, S., Simon, M. F., & Young, J. P. W. (2018). Horizontal transfer of symbiosis genes within and between rhizobial genera: occurrence and importance. **Genes** (Basel), 9(7). https://doi.org/10.3390/genes9070321

Andrio, E., Marino, D., Marmeys, A., de Segonzac, M. D., Damiani, I., Genre, A., Pauly, N. (2013). Hydrogen peroxide-regulated genes in the *Medicago truncatula-Sinorhizobium meliloti* symbiosis. **New Phytol,** 198(1), 179-189. https://doi.org/10.1111/nph.12120

Angot, A., Peeters, N., Lechner, E., Vailleau, F., Baud, C., Gentzbittel, L., Genin, S. (2006). *Ralstonia solanacearum* requires F-box-like domain-containing type III effectors to promote disease on several host plants. **Proc Natl Acad Sci U S A**, 103(39), 14620-14625. https://doi.org/10.1073/pnas.0509393103

Ardissone, S., Kobayashi, H., Kambara, K., Rummel, C., Noel, K. D., Walker, G. C., Deakin, W. J. (2011). Role of BacA in lipopolysaccharide synthesis, peptide transport, and nodulation by *Rhizobium* sp. strain NGR234. **J Bacteriol**, 193(9), 2218-2228. https://doi.org/10.1128/JB.01260-10

Arnold, M. F. F., Penterman, J., Shabab, M., Chen, E. J., & Walker, G. C. (2018). Important late-stage symbiotic role of the *Sinorhizobium meliloti* exopolysaccharide succinoglycan. **J Bacteriol**, 200(13). https://doi.org/10.1128/JB.00665-17

Arrighi, J. F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L. C., Mirabella, R., Gough, C. (2006). The *Medicago truncatula* lysin motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. **Plant Physiol**, 142(1), 265-279. https://doi.org/10.1104/pp.106.084657

Arrighi, J. F., Godfroy, O., de Billy, F., Saurat, O., Jauneau, A., & Gough, C. (2008). The RPG gene of *Medicago truncatula* controls *Rhizobium*-directed polar growth during infection. **Proc Natl Acad Sci U S A**, 105(28), 9817-9822. https://doi.org/10.1073/pnas.0710273105

Aslam, S. N., Newman, M. A., Erbs, G., Morrissey, K. L., Chinchilla, D., Boller, T., Cooper, R. M. (2008). Bacterial polysaccharides suppress induced innate immunity by calcium chelation. **Curr Biol**, 18(14), 1078-1083. https://doi.org/10.1016/j.cub.2008.06.061

Asolkar, T., & Ramesh, R. (2018). Development of T3SS mutants. **Indian J Microbiol**, 58(3), 372-380. https://doi.org/10.1007/s12088-018-0736-y

Barcellos, F. G., Menna, P., Batista, J. S. D., & Hungria, M. (2007). Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs *Sinorhizobium (Ensifer) fredii* and *Bradyrhizobium elkanii* in a

Brazilian savannah soil. **Applied and Environmental Microbiology**, 73(8), 2635-2643. https://doi.org/10.1128/aem.01823-06

Barnett, M. J., Fisher, R. F., Jones, T., Komp, C., Abola, A. P., Barloy-Hubler, F., Long, S. R. (2001). Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. **Proc Natl Acad Sci U S A**, 98(17), 9883-9888. https://doi.org/10.1073/pnas.161294798

Barrett, C. F., & Parker, M. A. (2006). Coexistence of *Burkholderia*, *Cupriavidus*, and *Rhizobium* sp. nodule bacteria on two *Mimosa* spp. in Costa Rica. **Appl Environ Microbiol**, 72(2), 1198-1206. https://doi.org/10.1128/AEM.72.2.1198-1206.2006

Bartsev, A. V., Deakin, W. J., Boukli, N. M., McAlvin, C. B., Stacey, G., Malnoe, P., Staehelin, C. (2004). NopL, an effector protein of *Rhizobium* sp NGR234, thwarts activation of plant defense reactions**.** **Plant Physiology**, 134(2), 871-879. https://doi.org/10.1104/pp.103.031740

Bataillon, T., & Bailey, S. F. (2014). Effects of new mutations on fitness: insights from models and data. **Ann N Y Acad Sci**, 1320, 76-92. https://doi.org/10.1111/nyas.12460

Batstone, R. T., O'Brien, A. M., Harrison, T. L., & Frederickson, M. E. (2020). Experimental evolution makes microbes more cooperative with their local host genotype. **Science**, 370(6515), 476-478. https://doi.org/10.1126/science.abb7222

Battisti, L., Lara, J. C., & Leigh, J. A. (1992). Specific oligosaccharide form of the *Rhizobium meliloti* exopolysaccharide promotes nodule invasion in alfalfa. **Proc Natl Acad Sci U S A**, 89(12), 5625-5629.

Bañuelos-Vazquez, L. A., Torres Tejerizo, G., Cervantes-De La Luz, L., Girard, L., Romero, D., & Brom, S. (2019). Conjugative transfer between *Rhizobium etli* endosymbionts inside the root nodule. **Environ Microbiol**. https://doi.org/10.1111/1462-2920.14645

Benezech, C., Doudement, M., & Gourion, B. (2020). Legumes tolerance to rhizobia is not always observed and not always deserved. **Cell Microbiol**, 22(1), e13124. https://doi.org/10.1111/cmi.13124

Berckx, F., Nguyen, T. V., Bandong, C. M., Lin, H. H., Yamanaka, T., Katayama, S., Pawlowski, K. (2022). A tale of two lineages: how the strains of the earliest divergent symbiotic Frankia clade spread over the world. **BMC Genomics**, 23(1), 602. https://doi.org/10.1186/s12864-022-08838-5

Bhagwat, A. A., Gross, K. C., Tully, R. E., & Keister, D. L. (1996). Beta-glucan synthesis in *Bradyrhizobium japonicum*: characterization of a new locus (*ndvC*) influencing beta-(1-->6) linkages. **J Bacteriol**, 178(15), 4635-4642.

Bhagwat, A. A., Tully, R. E., & Keister, D. L. (1992). Isolation and characterization of an *ndvB* locus from *Rhizobium fredii*. Mol Microbiol, 6(15), 2159-2165.

Black, M., Moolhuijzen, P., Chapman, B., Barrero, R., Howieson, J., Hungria, M., & Bellgard, M. (2012). The genetics of symbiotic nitrogen fixation: comparative genomics of 14 rhizobia strains by resolution of protein clusters. **Genes** (Basel), 3(1), 138-166. https://doi.org/10.3390/genes3010138

Bobik, C., Meilhoc, E., & Batut, J. (2006). FixJ: a major regulator of the oxygen limitation response and late symbiotic functions of *Sinorhizobium meliloti*. **Journal of Bacteriology**, 188(13), 4890-4902. https://doi.org/10.1128/jb.00251-06

Boivin, S., & Lepetit, M. (2020). Partner preference in the legume-rhizobia symbiosis and impact on legume inoculation strategies. In (Vol. 94 pp. 323–348). **Advances in Botanical Research: Regulation of Nitrogen-Fixing Symbioses in Legumes**.

Bontemps, C., Elliott, G. N., Simon, M. F., Dos Reis Junior, F. B., Gross, E., Lawton, R. C., Young, J. P. W. (2010). *Burkholderia* species are ancient symbionts of legumes. **Molecular Ecology**, 19(1), 44-52. https://doi.org/10.1111/j.1365-294X.2009.04458.x

Bontemps, C., Rogel, M. A., Wiechmann, A., Mussabekova, A., Moody, S., Simon, M. F., James, E. K. (2016). Endemic *Mimosa* species from Mexico prefer alphaproteobacterial rhizobial symbionts. **New Phytol**, 209(1), 319-333. https://doi.org/10.1111/nph.13573

Boucher, C. A., Barberis, P. A., Trigalet, A. P., & Demery, D. A. (1985). Transposon mutagenesis of *Pseudomonas solanacearum* - isolation of Tn5-induced avirulent mutants. **Journal of General Microbiology**, 131(SEP), 2449-2457.

Breedveld, M. W., & Miller, K. J. (1994). Cyclic beta-glucans of members of the family Rhizobiaceae. **Microbiol Rev**, 58(2), 145-161.

Brenner, A. E., Muñoz-Leal, S., Sachan, M., Labruna, M. B., & Raghavan, R. (2021). *Coxiella burnetii* and related tick endosymbionts evolved from pathogenic ancestors. **Genome Biol Evol**, 13(7). https://doi.org/10.1093/gbe/evab108

Broughton, W. J., Hanin, M., Relic, B., Kopciñska, J., Golinowski, W., Simsek, S., Deakin, W. J. (2006). Flavonoid-inducible modifications to rhamnan O antigens are necessary for *Rhizobium* sp. strain NGR234-legume symbioses. **J Bacteriol**, 188(10), 3654-3663. https://doi.org/10.1128/JB.188.10.3654-3663.2006

Burghardt, L. T., Guhlin, J., Chun, C. L., Liu, J., Sadowsky, M. J., Stupar, R. M., Tiffin, P. (2017). Transcriptomic basis of genome by genome variation in a legume-rhizobia mutualism. **Mol Ecol**, 26(21), 6122-6135. https://doi.org/10.1111/mec.14285

Caetano-Anollés, G., Wall, L., Micheli, A., Macchi, E., Bauer, W., & Favelukes, G. (1988). Role of motility and chemotaxis in efficiency of nodulation by *Rhizobium meliloti*. In (Vol. 86, pp. 1228-1235). **Plant Physiology**.

Campbell, G. R. O., Reuhs, B. L., & Walker, G. C. (2002). Chronic intracellular infection of alfalfa nodules by *Sinorhizobium meliloti* requires correct lipopolysaccharide core. **Proceedings of the National Academy of Sciences of the United States of America**, 99(6), 3938-3943. https://doi.org/10.1073/pnas.062425699

Capela, D., Filipe, C., Bobilk, C., Batut, J., & Bruand, C. (2006). *Sinorhizobium meliloti* differentiation during symbiosis with alfalfa: A transcriptomic dissection. **Molecular Plant-Microbe Interactions**, 19(4), 363-372. https://doi.org/10.1094/mpmi-19-0363

Capela, D., Marchetti, M., Clérissi, C., Perrier, A., Guetta, D., Gris, C., Masson-Boivin, C. (2017). Recruitment of a lineage-specific virulence regulatory pathway promotes intracellular infection by a plant pathogen experimentally evolved into a legume symbiont. **Mol Biol Evol**, 34(10), 2503-2521. https://doi.org/10.1093/molbev/msx165

Capoen, W., Goormachtig, S., & Holsters, M. (2010). Water-tolerant legume nodulation. **Journal of Experimental Botany**, 61(5), 1251-1255. https://doi.org/10.1093/jxb/erp326

Carroll, S. M., Chubiz, L. M., Agashe, D., & Marx, C. J. (2015). Parallel and divergent evolutionary solutions for the optimization of an engineered central metabolism in methylobacterium extorquens AM1. **Microorganisms**, 3(2), 152-174. https://doi.org/10.3390/microorganisms3020152

Cassman, K. G., & Dobermann, A. (2022). Nitrogen and the future of agriculture: 20 years on : This article belongs to Ambio's 50th Anniversary Collection. Theme: Solutions-oriented research. **Ambio**, 51(1), 17-24. https://doi.org/10.1007/s13280-021-01526-w

Catoira, R., Galera, C., de Billy, F., Penmetsa, R. V., Journet, E. P., Maillet, F., Dénarié, J. (2000). Four genes of *Medicago truncatula* controlling components of a nod factor transduction pathway. **Plant Cell**, 12(9), 1647-1666. https://doi.org/10.1105/tpc.12.9.1647

Cermola, M., Fedorova, E., Taté, R., Riccio, A., Favre, R., & Patriarca, E. J. (2000). Nodule invasion and symbiosome differentiation during *Rhizobium etli-Phaseolus vulgaris* symbiosis. **Mol Plant Microbe Interact**, 13(7), 733-741. https://doi.org/10.1094/MPMI.2000.13.7.733

Chaintreuil, C., Perrier, X., Martin, G., Fardoux, J., Lewis, G. P., Brottier, L., Arrighi, J. F. (2018). Naturally occurring variations in the nod-independent model legume *Aeschynomene evenia* and relatives: a resource for nodulation genetics. **BMC Plant Biol**, 18(1), 54. https://doi.org/10.1186/s12870-018-1260-2

Chen, W. F., Wang, E. T., Ji, Z. J., & Zhang, J. J. (2021). Recent development and new insight of diversification and symbiosis specificity of legume rhizobia: mechanism and application. **J Appl Microbiol**, 131(2), 553-563. https://doi.org/10.1111/jam.14960

Chen, W. M., James, E. K., Prescott, A. R., Kierans, M., & Sprent, J. I. (2003). Nodulation of *Mimosa* spp. by the beta-proteobacterium *Ralstonia taiwanensis*. **Molecular Plant-Microbe Interactions**, 16(12), 1051-1061.

Chen, W. M., Laevens, S., Lee, T. M., Coenye, T., De Vos, P., Mergeay, M., & Vandamme, P. (2001a). *Ralstonia taiwanensis* sp *nov*., isolated from root nodules of *Mimosa* species and sputum of a cystic fibrosis patient. **International Journal of Systematic and Evolutionary Microbiology**, 51, 1729-1735.

Chen, W. M., Laevens, S., Lee, T. M., Coenye, T., De Vos, P., Mergeay, M., & Vandamme, P. (2001b). *Ralstonia taiwanensis* sp. *nov*., isolated from root nodules of Mimosa species and sputum of a cystic fibrosis patient. **Int J Syst Evol Microbiol**, 51(Pt 5), 1729-1735. https://doi.org/10.1099/00207713-51-5-1729

Chen, X., & Zhang, J. (2014). Yeast mutation accumulation experiment supports elevated mutation rates at highly transcribed sites. **Proc Natl Acad Sci U S A**, 111(39), E4062. https://doi.org/10.1073/pnas.1412284111

Cheng, H. P., & Walker, G. C. (1998a). Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. **J Bacteriol**, 180(19), 5183-5191. https://doi.org/10.1128/JB.180.19.5183-5191.1998

Cheng, H. P., & Walker, G. C. (1998b). Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. **Journal of Bacteriology**, 180(19), 5183-5191.

Coba de la Peña, T., Fedorova, E., Pueyo, J. J., & Lucas, M. M. (2017). The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organelle? **Front Plant Sci**, 8, 2229. https://doi.org/10.3389/fpls.2017.02229

Cooper, V. S., Honsa, E., Rowe, H., Deitrick, C., Iverson, A. R., Whittall, J. J., Rosch, J. W. (2020). Experimental evolution. **mSystems**, 5(3). https://doi.org/10.1128/mSystems.00352-20

Corral, J., Sebastià, P., Coll, N. S., Barbé, J., Aranda, J., & Valls, M. (2020). Twitching and swimming motility play a role in *Ralstonia solanacearum* pathogenicity. **mSphere**, 5(2). https://doi.org/10.1128/mSphere.00740-19

Cosseau, C., & Batut, J. (2004). Genomics of the *ccoNOQP*-encoded cbb(3) oxidase complex in bacteria. **Archives of Microbiology**, 181(2), 89-96. https://doi.org/10.1007/s00203-003-0641-5

Cvijović, I., Nguyen Ba, A. N., & Desai, M. M. (2018). Experimental Studies of Evolutionary Dynamics in Microbes. **Trends Genet**, *34*(9), 693-703. https://doi.org/10.1016/j.tig.2018.06.004

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., Giraud, E. (2015). Convergent evolution of endosymbiont differentiation in Dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by Nodule-Specific Cysteine-Rich Peptides. **Plant Physiol**, 169(2), 1254-1265. https://doi.org/10.1104/pp.15.00584

D'Haeze, W., & Holsters, M. (2002). Nod factor structures, responses, and perception during initiation of nodule development. **Glycobiology**, 12(6), 79R-105R.

D'Haeze, W., & Holsters, M. (2004). Surface polysaccharides enable bacteria to evade plant immunity. **Trends in Microbiology**, 12(12), 555-561.

Dakora, F. D. (1995). A functional relationship between leghaemoglobin and nitrogenase based on novel measurements of the two proteins in legume root nodules. **Ann Bot**, 75(1), 49-54. https://doi.org/10.1016/S0305-7364(05)80008-3

Dalia, A. B., McDonough, E., & Camilli, A. (2014). Multiplex genome editing by natural transformation. **Proceedings of the National Academy of Sciences of the United States of America**, 111(24), 8937-8942. https://doi.org/10.1073/pnas.1406478111

Dalsing, B. L., & Allen, C. (2014). Nitrate assimilation contributes to *Ralstonia solanacearum* root attachment, stem colonization, and virulence. **J Bacteriol**, 196(5), 949-960. https://doi.org/10.1128/JB.01378-13

Dalsing, B. L., Truchon, A. N., Gonzalez-Orta, E. T., Milling, A. S., & Allen, C. (2015). *Ralstonia solanacearum* uses inorganic nitrogen metabolism for virulence, ATP production, and detoxification in the oxygen-limited host xylem environment. **mBio**, 6(2), e02471. https://doi.org/10.1128/mBio.02471-14

Daubech, B., Poinsot, V., Klonowska, A., Capela, D., Chaintreuil, C., Moulin, L., Masson-Boivin, C. (2019). A new nodulation gene involved in the biosynthesis of Nod Factors with an open-chain oxidized terminal residue and in the symbiosis with. **Mol Plant Microbe Interact**, 32(12), 1635-1648. https://doi.org/10.1094/MPMI-06-19-0168-R

Daubech, B., Remigi, P., Doin de Moura, G., Marchetti, M., Pouzet, C., Auriac, M. C., Capela, D. (2017). Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation. **Elife**, 6. https://doi.org/10.7554/eLife.28683

de Lajudie, P., & Young, J. P. W. (2019). International committee on systematics of Prokaryotes subcommittee on the taxonomy of rhizobia and agrobacteria. Minutes of the meeting by video conference, 11 July 2018. **Int J Syst Evol Microbiol**, 69(6), 1835-1840. https://doi.org/10.1099/ijsem.0.003335

de Lajudie, P. M., & Young, J. P. W. (2018). International committee on systematics of Prokaryotes subcommittee on the taxonomy of rhizobia and agrobacteria minutes of the

closed meeting, Granada, 4 September 2017. **Int J Syst Evol Microbiol**, 68(10), 3363-3368. https://doi.org/10.1099/ijsem.0.002974

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

Deatherage, D. E., & Barrick, J. E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. **Methods Mol Biol**, 1151, 165-188. https://doi.org/10.1007/978-1-4939-0554-6_12

Denamur, E., & Matic, I. (2006). Evolution of mutation rates in bacteria. **Molecular Microbiology**, 60(4), 820-827. https://doi.org/10.1111/j.1365-2958.2006.05150.x

Deslandes, L., Pileur, F., Liaubet, L., Camut, S., Can, C., Williams, K., Marco, Y. (1998). Genetic characterization of RRS1, a recessive locus in *Arabidopsis thaliana* that confers resistance to the bacterial soilborne pathogen *Ralstonia solanacearum*. **Mol Plant Microbe Interact**, 11(7), 659-667.

diCenzo, G. C., Zamani, M., Checcucci, A., Fondi, M., Griffitts, J. S., Finan, T. M., & Mengoni, A. (2019). Multidisciplinary approaches for studying *Rhizobium*-legume symbioses. **Can J Microbiol**, 65(1), 1-33. https://doi.org/10.1139/cjm-2018-0377

diCenzo, G. C., Zamani, M., Cowie, A., & Finan, T. M. (2015). Proline auxotrophy in *Sinorhizobium meliloti* results in a plant-specific symbiotic phenotype. **Microbiology**, 161(12), 2341-2351. https://doi.org/10.1099/mic.0.000182

Ding, H., Yip, C. B., Geddes, B. A., Oresnik, I. J., & Hynes, M. F. (2012). Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. **Microbiology**, 158(Pt 5), 1369-1378. https://doi.org/10.1099/mic.0.057281-0

Doin de Moura, G. G., Mouffok, S., Gaudu, N., Cazalé, A.-C., Milhes, M., Bulach, T., Remigi, P. V. O. P. (2022). A selective bottleneck during host entry drives the evolution of new legume symbionts. In (pp. 1-35). **BioRxiv**.

Doin de Moura, G. G., Remigi, P., Masson-Boivin, C., & Capela, D. (2020). Experimental Evolution of Legume symbionts: what have we learnt? **Genes** (Basel), 11(3). https://doi.org/10.3390/genes11030339

Dong, W., & Song, Y. (2020). The significance of flavonoids in the process of biological nitrogen fixation. **Int J Mol Sci**, 21(16). https://doi.org/10.3390/ijms21165926

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

Downie, J. A. (2014). Legume nodulation. **Curr Biol**, 24(5), R184-190. https://doi.org/10.1016/j.cub.2014.01.028

Drew, G. C., Stevens, E. J., & King, K. C. (2021). Microbial evolution and transitions along the parasite-mutualist continuum. **Nat Rev Microbiol**, 19(10), 623-638. https://doi.org/10.1038/s41579-021-00550-7

Dunn, M. F. (2014). Key roles of microsymbiont amino acid metabolism in rhizobia-legume interactions. In (Vol. 41, pp. 411–451). **Critical Reviews in Microbiology**.

Dunn, M. F. (2015). Key roles of microsymbiont amino acid metabolism in rhizobia-legume interactions. **Crit Rev Microbiol**, 41(4), 411-451. https://doi.org/10.3109/1040841X.2013.856854

Dylan, T., Nagpal, P., Helinski, D. R., & Ditta, G. S. (1990). Symbiotic pseudorevertants of *Rhizobium meliloti ndv* mutants. **J Bacteriol**, 172(3), 1409-1417.

Dymov, S. I., Meek, D. J., Steven, B., & Driscoll, B. T. (2004). Insertion of transposon Tn5tac1 in the *Sinorhizobium meliloti* malate dehydrogenase (mdh) gene results in conditional polar effects on downstream TCA cycle genes. **Mol Plant Microbe Interact**, 17(12), 1318-1327. https://doi.org/10.1094/MPMI.2004.17.12.1318

Elliott, G. N., Chou, J. H., Chen, W. M., Bloemberg, G. V., Bontemps, C., Martinez-Romero, E., James, E. K. (2009). *Burkholderia* spp. are the most competitive symbionts of *Mimosa*, particularly under N-limited conditions. **Environmental Microbiology**, 11(4), 762-778. https://doi.org/10.1111/j.1462-2920.2008.01799.x

Erhardt, M., Namba, K., & Hughes, K. T. (2010). Bacterial nanomachines: the flagellum and type III injectisome. **Cold Spring Harb Perspect Biol**, 2(11), a000299. https://doi.org/10.1101/cshperspect.a000299

Fabre, S., Gully, D., Poitout, A., Patrel, D., Arrighi, J. F., Giraud, E., Cartieaux, F. (2015). Nod Factor-Independent nodulation in *Aeschynomene evenia* Required the Common Plant-Microbe Symbiotic Toolkit. **Plant Physiol**, 169(4), 2654-2664. https://doi.org/10.1104/pp.15.01134

Faget, M., Blossfeld, S., von Gillhaussen, P., Schurr, U., & Temperton, V. M. (2013). Disentangling who is who during rhizosphere acidification in root interactions: combining fluorescence with optode techniques. **Front Plant Sci**, 4, 392. https://doi.org/10.3389/fpls.2013.00392

Fagorzi, C., Ilie, A., Decorosi, F., Cangioli, L., Viti, C., Mengoni, A., & diCenzo, G. C. (2020). Symbiotic and nonsymbiotic members of the genus *Ensifer* (syn. *Sinorhizobium*) are separated into two clades based on comparative genomics and high-throughput phenotyping. **Genome Biol Evol**, 12(12), 2521-2534. https://doi.org/10.1093/gbe/evaa221

FAO. (2019). The State of Food and Agriculture 2019. Moving forward on food loss and waste reduction.

Farkas, A., Maróti, G., Durgő, H., Györgypál, Z., Lima, R. M., Medzihradszky, K. F., Kondorosi, É. (2014). *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. **Proc Natl Acad Sci U S A**, 111(14), 5183-5188. https://doi.org/10.1073/pnas.1404169111

Fischer, H. M. (1994). Genetic regulation of nitrogen fixation in rhizobia. **Microbiol Rev**, 58(3), 352-386. https://doi.org/10.1128/mr.58.3.352-386.1994

Flores-Tinoco, C. E., Tschan, F., Fuhrer, T., Margot, C., Sauer, U., Christen, M., & Christen, B. (2020). Co-catabolism of arginine and succinate drives symbiotic nitrogen fixation. **Mol Syst Biol**, 16(6), e9419. https://doi.org/10.15252/msb.20199419

Flórez, L. V., Scherlach, K., Gaube, P., Ross, C., Sitte, E., Hermes, C., Kaltenpoth, M. (2017). Antibiotic-producing symbionts dynamically transition between plant pathogenicity and insect-defensive mutualism. **Nat Commun**, 8, 15172. https://doi.org/10.1038/ncomms15172

Foster, P. L. (2007). Stress-induced mutagenesis in bacteria. **Critical Reviews in Biochemistry and Molecular Biology**, 42(5), 373-397. https://doi.org/10.1080/10409230701648494

Fougère, F., & Le Rudulier, D. (1990). Uptake of glycine betaine and its analogues by bacteroids of *Rhizobium meliloti*. **J Gen Microbiol**, 136(1), 157-163. https://doi.org/10.1099/00221287-136-1-157

Fraysse, N., Couderc, F., & Poinsot, V. (2003). Surface polysaccharide involvement in establishing the rhizobium-legume symbiosis. **European Journal of Biochemistry**, 270(7), 1365-1380. https://doi.org/10.1046/j.1432-1033.2003.03492.x

Friedman, L., & Kolter, R. (2004). Two genetic loci produce distinct carbohydrate-rich structural components of the *Pseudomonas aeruginosa* biofilm matrix. **J Bacteriol**, 186(14), 4457-4465. https://doi.org/10.1128/JB.186.14.4457-4465.2004

Friesen, M. L. (2012). Widespread fitness alignment in the legume-rhizobium symbiosis. **New Phytologist**, 194(4), 1096-1111. https://doi.org/10.1111/j.1469-8137.2012.04099.x

Fujihara, S., & Yoneyama, T. (1993). Effects of pH and osmotic stress on cellular polyamine contents in the soybean rhizobia *Rhizobium fredii* P220 and *Bradyrhizobium japonicum* A1017. **Appl Environ Microbiol**, 59(4), 1104-1109. https://doi.org/10.1128/aem.59.4.1104-1109.1993

Gage, D. J. (2002). Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*. **Journal of Bacteriology**, 184(24), 7042-7046. https://doi.org/10.1128/jb.184.24.7042-7046.2002

Gage, D. J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. **Microbiology and Molecular Biology Reviews**, 68(2), 280-+. https://doi.org/10.1128/mmbr.68.2.280-300.2004

Galhardo, R. S., Hastings, P. J., & Rosenberg, S. M. (2007). Mutation as a stress response and the regulation of evolvability. **Critical Reviews in Biochemistry and Molecular Biology**, 42(5), 399-435. https://doi.org/10.1080/10409230701648502

Galibert, F., Finan, T. M., Long, S. R., Puhler, A., Abola, P., Ampe, F., Batut, J. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. **Science**, 293(5530), 668-672.

Galindo, F. S., Pagliari, P. H., da Silva, E. C., Silva, V. M., Fernandes, G. C., Rodrigues, W. L., Teixeira Filho, M. C. M. (2022). Co-Inoculation with. Plants (Basel), 11(14). https://doi.org/10.3390/plants11141847

Garrido-Oter, R., Nakano, R. T., Dombrowski, N., Ma, K. W., McHardy, A. C., Schulze-Lefert, P., & Team, A. (2018). Modular traits of the rhizobiales root microbiota and their evolutionary relationship with symbiotic rhizobia. **Cell Host Microbe**, 24(1), 155-167.e155. https://doi.org/10.1016/j.chom.2018.06.006

Geddes, B. A., González, J. E., & Oresnik, I. J. (2014). Exopolysaccharide production in response to medium acidification is correlated with an increase in competition for nodule occupancy. **Mol Plant Microbe Interact**, 27(12), 1307-1317. https://doi.org/10.1094/MPMI-06-14-0168-R

Geddes, B. A., Kearsley, J. V. S., Huang, J., Zamani, M., Muhammed, Z., Sather, L., Finan, T. M. (2021). Minimal gene set from. **Proc Natl Acad Sci U S A**, 118(2). https://doi.org/10.1073/pnas.2018015118

Gehlot, tak, & gyaneshwar, a. (2013). An invasive *Mimosa* in India does not adopt the symbiont of its native relatives. In: **Annals of botany**.

Genin, S., & Denny, T. P. (2012). Pathogenomics of the *Ralstonia solanacearum* species complex. **Annual Review of Phytopathology**, Vol 50, 50, 67-89. https://doi.org/10.1146/annurev-phyto-081211-173000

Giraud, E., Hannibal, L., Fardoux, L., Vermeglio, 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.

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J. C., Sadowsky, M. (2007). Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. **Science**, 316(5829), 1307-1312. https://doi.org/10.1126/science.1139548

González, A. H., Morales Londoño, D., Pille da Silva, E., Nascimento, F. X. I., de Souza, L. F., da Silva, B. G., Soares, C. R. F. S. (2019). *Bradyrhizobium* and *Pseudomonas* strains obtained from coal-mining areas nodulate and promote the growth of *Calopogonium muconoides* plants used in the reclamation of degraded areas. **J Appl Microbiol**, 126(2), 523-533. https://doi.org/10.1111/jam.14117

Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., & Desai, M. M. (2017). The dynamics of molecular evolution over 60,000 generations. **Nature**, 551(7678), 45-50. https://doi.org/10.1038/nature24287

Gossmann, J. A., Markmann, K., Brachmann, A., Rose, L. E., & Parniske, M. (2012). Polymorphic infection and organogenesis patterns induced by a *Rhizobium leguminosarum* isolate from *Lotus* root nodules are determined by the host genotype. **New Phytol**, 196(2), 561-573. https://doi.org/10.1111/j.1469-8137.2012.04281.x

Gostinčar, C., Stajich, J. E., Kejžar, A., Sinha, S., Nislow, C., Lenassi, M., & Gunde-Cimerman, N. (2021). Seven years at high salinity-experimental evolution of the extremely halotolerant black yeast. **J Fungi** (Basel), 7(9). https://doi.org/10.3390/jof7090723

Gough, C., & Cullimore, J. (2011). Lipo-chitooligosaccharide Signaling in endosymbiotic plant-microbe Interactions. **Molecular Plant-Microbe Interactions**, 24(8), 867-878. https://doi.org/10.1094/mpmi-01-11-0019

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

Green, R. T., East, A. K., Karunakaran, R., Downie, J. A., & Poole, P. (2019). Transcriptomic analysis of *Rhizobium leguminosarum* bacteroids in determinate and indeterminate nodules. In (Vol. 5, pp. 1-16). **Microbial Genomics**.

Green, R. T., East, A. K., Karunakaran, R., Downie, J. A., & Poole, P. S. (2019). Transcriptomic analysis of *Rhizobium leguminosarum* bacteroids in determinate and indeterminate nodules. **Microb Genom**, 5(2). https://doi.org/10.1099/mgen.0.000254

Griesmann, M., Chang, Y., Liu, X., Song, Y., Haberer, G., Crook, M. B., Cheng, S. (2018). Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. **Science**, 361(6398). https://doi.org/10.1126/science.aat1743

Guan, N., & Liu, L. (2020). Microbial response to acid stress: mechanisms and applications. **Appl Microbiol Biotechnol**, 104(1), 51-65. https://doi.org/10.1007/s00253-019-10226-1

Guan, S. H., Gris, C., Cruveiller, S., Pouzet, C., Tasse, L., Leru, A., Capela, D. (2013). Experimental evolution of nodule intracellular infection in legume symbionts. **Isme Journal**, 7(7), 1367-1377. https://doi.org/10.1038/ismej.2013.24

Gubry-Rangin, C., Garcia, M., & Bena, G. (2010). Partner choice in *Medicago truncatula-Sinorhizobium* symbiosis. **Proceedings of the Royal Society B-Biological Sciences**, 277(1690), 1947-1951. https://doi.org/10.1098/rspb.2009.2072

Guefrachi, I., Pierre, O., Timchenko, T., Alunni, B., Barrière, Q., Czernic, P., Mergaert, P. (2015). *Bradyrhizobium* BclA is a peptide transporter required for bacterial differentiation in symbiosis with *Aeschynomene* legumes. **Mol Plant Microbe Interact**, 28(11), 1155-1166. https://doi.org/10.1094/MPMI-04-15-0094-R

Guidot, A., Jiang, W., Ferdy, J.-B., Thebaud, C., Barberis, P., Gouzy, J., & Genin, S. (2014). Multihost experimental evolution of the Pathogen *Ralstonia solanacearum* unveils genes involved in adaptation to plants. **Molecular Biology and Evolution**, 31(11), 2913-2928. https://doi.org/10.1093/molbev/msu229

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. **Syst Biol**, 59(3), 307-321. https://doi.org/10.1093/sysbio/syq010

Gully, D., Czernic, P., Cruveiller, S., Mahé, F., Longin, C., Vallenet, D., Cartieaux, F. (2018). Transcriptome profiles of Nod Factor-independent symbiosis in the tropical legume *Aeschynomene evenia***. Sci Rep**, 8(1), 10934. https://doi.org/10.1038/s41598-018-29301-0

Haag, A. F., Baloban, M., Sani, M., Kerscher, B., Pierre, O., Farkas, A., Ferguson, G. P. (2011). Protection of *Sinorhizobium* against host cysteine-rich antimicrobial peptides is critical for symbiosis. **PLoS Biol**, 9(10), e1001169. https://doi.org/10.1371/journal.pbio.1001169

Haag , A. F., & Mergaert , P. (2019). Terminal bacteroid differentiation in the *Medicago–Rhizobium* interaction – a tug of war between plant and bacteria. In. **Wiley online library**.

Hahn, M., & Studer, D. (1986). Competitiveness of a *nif− Bradyrhizobium japonicum* mutant against the wild-type strain. **FEMS Microbiology Letters**, Vol. 33, pp. 143-148.

Haskett, T. L., Terpolilli, J. J., Bekuma, A., O'Hara, G. W., Sullivan, J. T., Wang, P., Ramsay, J. P. (2016). Assembly and transfer of tripartite integrative and conjugative genetic elements. **Proc Natl Acad Sci U S A**, 113(43), 12268-12273. https://doi.org/10.1073/pnas.1613358113

Haskett, T. L., Terpolilli, J. J., Ramachandran, V. K., Verdonk, C. J., Poole, P. S., O'Hara, G. W., & Ramsay, J. P. (2018). Sequential induction of three recombination directionality factors directs assembly of tripartite integrative and conjugative elements. **PLoS Genet**, 14(3), e1007292. https://doi.org/10.1371/journal.pgen.1007292

Hassan, S., & Mathesius, U. (2012). The role of flavonoids in root-rhizosphere signalling: opportunities and challenges for improving plant-microbe interactions. **J Exp Bot**, 63(9), 3429-3444. https://doi.org/10.1093/jxb/err430

Hausmann, S., Gonzalez, D., Geiser, J., & Valentini, M. (2021). The DEAD-box RNA helicase RhlE2 is a global regulator of *Pseudomonas aeruginosa* lifestyle and pathogenesis. **Nucleic Acids Res**, 49(12), 6925-6940. https://doi.org/10.1093/nar/gkab503

Heath, K. D., & Tiffin, P. (2009). Stabilizing mechanisms in a legume-rhizobium mutualism. **Evolution**, 63(3), 652-662. https://doi.org/10.1111/j.1558-5646.2008.00582.x

Hinde, R., & Trautman, D. A. (2001). Symbiosomes. In (Vol. 4). **Symbiosis. Cellular Origin, Life in Extreme Habitats and Astrobiology**: Springer.

Hindré, T., Knibbe, C., Beslon, G., & Schneider, D. (2012). New insights into bacterial adaptation through *in vivo* and *in silico* experimental evolution. **Nat Rev Microbiol**, 10(5), 352-365. https://doi.org/10.1038/nrmicro2750

Hirsch, A. M., Wilson, K. J., Jones, J. D. G., Bang, M., Walker, V. V., & Ausubel, F. M. (1984). *Rhizobium meliloti* nodulation genes allow *Agrobacterium tumefasciens* and *Escherichia coli* to form pseudonodules on alfafa. **Journal of Bacteriology**, 158(3), 1133-1143.

Hosokawa, T., Ishii, Y., Nikoh, N., Fujie, M., Satoh, N., & Fukatsu, T. (2016). Obligate bacterial mutualists evolving from environmental bacteria in natural insect populations. **Nat Microbiol**, 1, 15011. https://doi.org/10.1038/nmicrobiol.2015.11

Huang, C. J., Lu, M. Y., Chang, Y. W., & Li, W. H. (2018). Experimental evolution of yeast for high-temperature tolerance. **Mol Biol Evol**, 35(8), 1823-1839. https://doi.org/10.1093/molbev/msy077

Hubber, A., Vergunst, A. C., Sullivan, J. T., Hooykaas, P. J. J., & Ronson, C. W. (2004). Symbiotic phenotypes and translocated effector proteins of the *Mesorhizobium loti* strain R7A VirB/D4 type IV secretion system. **Molecular Microbiology**, 54(2), 561-574. https://doi.org/10.1111/j.1365-2958.2004.04292.x

Hug, S., Liu, Y., Heiniger, B., Bailly, A., Ahrens, C. H., Eberl, L., & Pessi, G. (2021). Differential Expression of. **Front Plant Sci**, 12, 699590. https://doi.org/10.3389/fpls.2021.699590

Hussain , M. Z., Bhardwaj , A. K., Basso , G. B., Robertson , P., & Hamilton , S. K. (2019). Nitrate leaching from continuous corn, perennial grasses, and poplar in the US Midwest. **Journal of Environmental Quality**.

Imlay, J. A. (2013). The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. **Nature Reviews Microbiology**, 11(7), 443-454. https://doi.org/10.1038/nrmicro3032

Inoue, T., Higuchi, M., Hashimoto, Y., Seki, M., Kobayashi, M., Kato, T., Kakimoto, T. (2001). Identification of CRE1 as a cytokinin receptor from *Arabidopsis*. **Nature**, 409(6823), 1060-1063. https://doi.org/10.1038/35059117

Jamet, A., Mandon, K., Puppo, A., & Hérouart, D. (2007). H_2O_2 is required for optimal establishment of the *Medicago sativa*/*Sinorhizobium meliloti* symbiosis. **J Bacteriol**, 189(23), 8741-8745. https://doi.org/10.1128/JB.01130-07

Jaszek, M., Janczarek, M., Piersiak, T., Grzywnowicz, & K. (2014). The response of the *Rhizobium leguminosarum* bv. *trifolii* wild-type and exopolysaccharide-deficient mutants to oxidative stress. **Plant soil**, Vol. 376, pp. 75–94.

Jeudy, C., Ruffel, S., Freixes, S., Tillard, P., Santoni, A. L., Morel, S., Salon, C. (2010). Adaptation of *Medicago truncatula* to nitrogen limitation is modulated via local and systemic nodule developmental responses. **New Phytol**, 185(3), 817-828. https://doi.org/10.1111/j.1469-8137.2009.03103.x

Jiménez-Zurdo, J. I., van Dillewijn, P., Soto, M. J., de Felipe, M. R., Olivares, J., & Toro, N. (1995). Characterization of a *Rhizobium meliloti* proline dehydrogenase mutant altered in nodulation efficiency and competitiveness on alfalfa roots. **Mol Plant Microbe Interact**, 8(4), 492-498. https://doi.org/10.1094/mpmi-8-0492

Jiménez-Zurdo, J. I., van Dillewijn, P., Soto, M. J., de Felipe, M. R., Olivares, J., & Toro, N. (1995). Characterization of a *Rhizobium meliloti* proline dehydrogenase mutant altered in nodulation efficiency and competitiveness on alfalfa roots. **Molecular Plant-Microbe Interaction**, Vol. 8, pp. 492/498..

Jones, D. L., Dennis, P. G., Owen, A. G., & van Hees, P. A. W. (2003). Organic acid behavior in soils - misconceptions and knowledge gaps. **Plant and Soil**, vol. 248, pp. 31-41.

Jones, D. L., Farrar, J., & Giller, K. E. (2003). Associative nitrogen fixation and root exudation - what is theoretically possible in the rhizosphere? **Symbiosis**, v. 35.

Jones, K. M., Sharopova, N., Lohar, D. P., Zhang, J. Q., VandenBosch, K. A., & Walker, G. C. (2008). Differential response of the plant *Medicago truncatula* to its symbiont *Sinorhizobium meliloti* or an exopolysaccharide-deficient mutant. **Proceedings of the National Academy of Sciences of the United States of America**, 105(2), 704-709. https://doi.org/10.1073/pnas.0709338105

Jourand, P., Renier, A., Rapior, S., de Faria, S. M., Prin, Y., Galiana, A., Dreyfus, B. (2005). Role of methylotrophy during symbiosis between *Methylobacterium nodulans* and *Crotalaria podocarpa*. **Molecular Plant-Microbe Interactions**, 18(10), 1061-1068. https://doi.org/10.1094/mpmi-18-1061

Kang, Y. H., J. Mao, G. He, L.Y. (1994). Dramatically reduced virulence of mutants of *Pseudomonas solanacearum* Defective in export of extracellular proteins across the outer membrane. **Molecular Plant-Microbe Interactions**, v. 7.

Karunakaran, R., Haag, A. F., East, A. K., Ramachandran, V. K., Prell, J., James, E. K., Poole, P. S. (2010). BacA is essential for bacteroid development in nodules of Galegoid, but not Phaseoloid, legumes. **Journal of Bacteriology**, 192(11), 2920-2928. https://doi.org/10.1128/jb.00020-10

Kassen, R., & Bataillon, T. (2006). Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. **Nat Genet**, 38(4), 484-488. https://doi.org/10.1038/ng1751

Kawaharada, Y., Kelly, S., Nielsen, M. W., Hjuler, C. T., Gysel, K., Muszyński, A., Stougaard, J. (2015). Receptor-mediated exopolysaccharide perception controls bacterial infection. **Nature**, 523(7560), 308-312. https://doi.org/10.1038/nature14611

Kawaharada, Y., Nielsen, M. W., Kelly, S., James, E. K., Andersen, K. R., Rasmussen, S. R., Stougaard, J. (2017). Differential regulation of the Epr3 receptor coordinates membrane-restricted rhizobial colonization of root nodule primordia. **Nat Commun**, 8, 14534. https://doi.org/10.1038/ncomms14534

Khokhani, D., Lowe-Power, T. M., Tran, T. M., & Allen, C. (2017). A single regulator mediates strategic switching between attachment/spread and growth/virulence in the plant pathogen. **MBio**, 8(5). https://doi.org/10.1128/mBio.00895-17

Kiers, E., Rousseau, R., West, S., & Denison, R. (2003). Host sanctions and the legume-rhizobium mutualism. **Nature**, 425(6953), 78-81.

Kiers, E. T., & Denison, R. F. (2008). Sanctions, cooperation, and the stability of plant-rhizosphere mutualisms. **Annual Review of Ecology Evolution and Systematics**, 39, 215-236. https://doi.org/10.1146/annurev.ecolsys.39.110707.173423

Kiers, E. T., Ratcliff, W. C., & Denison, R. F. (2013). Single-strain inoculation may create spurious correlations between legume fitness and rhizobial fitness. **New Phytologist**, 198(1), 4-6. https://doi.org/10.1111/nph.12015

Kim, C. H., Tully, R. E., & Keister, D. L. (1989). Exopolysaccharide-deficient mutants of *Rhizobium fredii* HH303 which are symbiotically effective. **Appl Environ Microbiol**, 55(7), 1852-1854.

King, K. C., Brockhurst, M. A., Vasieva, O., Paterson, S., Betts, A., Ford, S. A., Hurst, G. D. (2016). Rapid evolution of microbe-mediated protection against pathogens in a worm host. **ISME J,** 10(8), 1915-1924. https://doi.org/10.1038/ismej.2015.259

Klonowska, A., Chaintreuil, C., Tisseyre, P., Miche, L., Melkonian, R., Ducousso, M., Moulin, L. (2012). Biodiversity of *Mimosa pudica* rhizobial symbionts (*Cupriavidus taiwanensis*, *Rhizobium mesoamericanum*) in New Caledonia and their adaptation to heavy metal-rich soils. **Fems Microbiology Ecology**, 81(3), 618-635. https://doi.org/10.1111/j.1574-6941.2012.01393.x

Koga, R., Moriyama, M., Onodera-Tanifuji, N., Ishii, Y., Takai, H., Mizutani, M., Fukatsu, T. (2022). Single mutation makes *Escherichia coli* an insect mutualist. **Nat Microbiol**, 7(8), 1141-1150. https://doi.org/10.1038/s41564-022-01179-9

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

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. **Proc Natl Acad Sci U S A**, 84(21), 7428-7432. https://doi.org/10.1073/pnas.84.21.7428

Kouchi, H., Shimomura, K., Hata, S., Hirota, A., Wu, G. J., Kumagai, H., Tabata, S. (2004). Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*. **DNA Res**, 11(4), 263-274. https://doi.org/10.1093/dnares/11.4.263

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), 1228-1235. https://doi.org/10.1094/MPMI.2002.15.12.1228

Kumar, A., Vij, N., & Randhawa, G. S. (2003). Isolation and symbiotic characterization of transposon Tn5-induced arginine auxotrophs of *Sinorhizobium meliloti***. Indian J Exp Biol**, 41(10), 1198-1204.

Köhler, T., Curty, L. K., Barja, F., van Delden, C., & Pechère, J. C. (2000). Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires flagella and

pili. **J Bacteriol**, 182(21), 5990-5996. https://doi.org/10.1128/JB.182.21.5990-5996.2000

Ladha, J. K., Tirol-Padre, A., Reddy, C. K., Cassman, K. G., Verma, S., Powlson, D. S., Pathak, H. (2016). Global nitrogen budgets in cereals: A 50-year assessment for maize, rice, and wheat production systems. **Sci Rep**, 6, 19355. https://doi.org/10.1038/srep19355

Laffont, C., Ivanovici, A., Gautrat, P., Brault, M., Djordjevic, M. A., & Frugier, F. (2020). The NIN transcription factor coordinates CEP and CLE signaling peptides that regulate nodulation antagonistically. **Nat Commun**, 11(1), 3167. https://doi.org/10.1038/s41467-020-16968-1

Laguerre, G., Heulin-Gotty, K., Brunel, B., Klonowska, A., Le Quere, A., Tillard, P., Lepetit, M. (2012). Local and systemic N signaling are involved in *Medicago truncatula* preference for the most efficient *Sinorhizobium symbiotic* partners. **New Phytologist**, 195(2), 437-449. https://doi.org/10.1111/j.1469-8137.2012.04159.x

Lam, H., Oh, D. C., Cava, F., Takacs, C. N., Clardy, J., de Pedro, M. A., & Waldor, M. K. (2009). D-amino acids govern stationary phase cell wall remodeling in bacteria. **Science,** 325(5947), 1552-1555. https://doi.org/10.1126/science.1178123

Lang, G. I., Rice, D. P., Hickman, M. J., Sodergren, E., Weinstock, G. M., Botstein, D., & Desai, M. M. (2013). Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. **Nature,** 500(7464), 571-574. https://doi.org/10.1038/nature12344

Laporte, P., Lepage, A., Fournier, J., Catrice, O., Moreau, S., Jardinaud, M. F., Niebel, A. (2014). The CCAAT box-binding transcription factor NF-YA1 controls rhizobial infection. **J Exp Bot**, 65(2), 481-494. https://doi.org/10.1093/jxb/ert392

Larsen, K., & Jochimsen, B. U. (1987). Appearance of purine-catabolizing enzymes in fix and fix root nodules on soybean and effect of oxygen on the expression of the enzymes in callus tissue. **Plant Physiol,** 85(2), 452-456. https://doi.org/10.1104/pp.85.2.452

Le, S. Q., & Gascuel, O. (2008). An improved general amino acid replacement matrix. **Mol Biol Evol**, 25(7), 1307-1320. https://doi.org/10.1093/molbev/msn067

Lebedeva, M., Azarakhsh, M., Sadikova, D., & Lutova, L. (2021). At the root of nodule organogenesis: conserved regulatory pathways recruited by rhizobia. **Plants** (Basel), 10(12). https://doi.org/10.3390/plants10122654

Ledbetter, R. N., Garcia Costas, A. M., Lubner, C. E., Mulder, D. W., Tokmina-Lukaszewska, M., Artz, J. H., Seefeldt, L. C. (2017). The electron bifurcating FixABCX protein complex from *Azotobacter vinelandii:* generation of low-potential reducing equivalents for nitrogenase catalysis. **Biochemistry,** 56(32), 4177-4190. https://doi.org/10.1021/acs.biochem.7b00389

Ledermann, R., Schulte, C. C. M., & Poole, P. S. (2021). How rhizobia adapt to the nodule environment. **J Bacteriol**, 203(12), e0053920. https://doi.org/10.1128/JB.00539-20

Lei, N., Chen, L., Kiba, A., Hikichi, Y., Zhang, Y., & Ohnishi, K. (2020). Super-multiple deletion analysis of Type III Effectors in. **Front Microbiol**, 11, 1683. https://doi.org/10.3389/fmicb.2020.01683

Lemaire, B., Van Cauwenberghe, J., Chimphango, S., Stirton, C., Honnay, O., Smets, E., & Muasya, A. M. (2015). Recombination and horizontal transfer of nodulation and ACC deaminase (*acdS)* genes within Alpha- and Betaproteobacteria nodulating legumes of

the Cape Fynbos biome. **FEMS Microbiol Ecol**, 91(11). https://doi.org/10.1093/femsec/fiv118

Lenski, R. E., Rose, M. R., Simpson, S. C., & Tadler, S. C. (1991). Long-term experimental evolution in *Escherichia coli*. I. adaptation and divergence during 2,000 generations. In (Vol. 138). **The American Naturalist.**

Lenski, R. E., & Travisano, M. (1994).Dynamics of adaptation and diversification – a 10000 generation experiment with bacterial populations. **Proceedings of the National Academy of Sciences of the United States of America**, 91(15), 6808-6814. https://doi.org/10.1073/pnas.91.15.6808

Lescat, M., Launay, A., Ghalayini, M., Magnan, M., Glodt, J., Pintard, C., Tenaillon, O. (2017). Using long-term experimental evolution to uncover the patterns and determinants of molecular evolution of an *Escherichia coli* natural isolate in the streptomycin-treated mouse gut. **Mol Ecol**, 26(7), 1802-1817. https://doi.org/10.1111/mec.13851

Li, E., de Jonge, R., Liu, C., Jiang, H., Friman, V. P., Pieterse, C. M. J., Jousset, A. (2021). Rapid evolution of bacterial mutualism in the plant rhizosphere. **Nat Commun**, 12(1), 3829. https://doi.org/10.1038/s41467-021-24005-y

Liang, Y., Cao, Y. R., Tanaka, K., Thibivilliers, S., Wan, J. R., Choi, J., Stacey, G. (2013). Nonlegumes respond to rhizobial Nod Factors by suppressing the innate immune response. **Science**, 341(6152), 1384-1387. https://doi.org/10.1126/science.1242736

Libault, M., Farmer, A., Brechenmacher, L., Drnevich, J., Langley, R. J., Bilgin, D. 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 Physiol**, 152(2), 541-552. https://doi.org/10.1104/pp.109.148379

Libourel , V. O. P. C., Keller , J. V. O. P., Brichet , L., Cazalé , A.-C., Carrère , S., Vernié , T., Capela , D. (2022). Comparative phylotranscriptomics reveals a 110 million years-old symbiotic program. In. **bioRvix**.

Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., & Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. **Science**, 302(5645), 630-633. https://doi.org/10.1126/science.1090074

Ling, J., Wang, H., Wu, P., Li, T., Tang, Y., Naseer, N., Zhu, J. (2016). Plant nodulation inducers enhance horizontal gene transfer of *Azorhizobium caulinodans* symbiosis island. **Proc Natl Acad Sci U S A**, 113(48), 13875-13880. https://doi.org/10.1073/pnas.1615121113

Liu, C. W., Breakspear, A., Guan, D., Cerri, M. R., Jackson, K., Jiang, S., Murray, J. D. (2019). NIN acts as a network hub controlling a growth module required for rhizobial Infection. **Plant Physiol**, 179(4), 1704-1722. https://doi.org/10.1104/pp.18.01572

Liu, J., You, L., Amini, M., Obersteiner, M., Herrero, M., Zehnder, A. J., & Yang, H. (2010). A high-resolution assessment on global nitrogen flows in cropland. **Proc Natl Acad Sci U S A**, 107(17), 8035-8040. https://doi.org/10.1073/pnas.0913658107

Liu, X. Y., Wu, W., Wang, E. T., Zhang, B., Macdermott, J., & Chen, W. X. (2011). Phylogenetic relationships and diversity of β-rhizobia associated with *Mimosa* species grown in Sishuangbanna, China. **Int J Syst Evol Microbiol**, 61(Pt 2), 334-342. https://doi.org/10.1099/ijs.0.020560-0

Liu, Y., Jiang, X., Guan, D., Zhou, W., Ma, M., Zhao, B., Li, J. (2017). Transcriptional analysis of genes involved in competitive nodulation in Bradyrhizobium diazoefficiens at the

presence of soybean root exudates. **Sci Rep**, 7(1), 10946. https://doi.org/10.1038/s41598-017-11372-0

Liu, Y., Tan, X., Pan, Y., Yu, J., Du, Y., Liu, X., & Ding, W. (2022). Mutation in *phcA* enhanced the adaptation of *Ralstonia solanacearum* to long term acid stress. **Front Microbiol**, 13, 829719. https://doi.org/10.3389/fmicb.2022.829719

Lodwig, E. M., Leonard, M., Marroqui, S., Wheeler, T. R., Findlay, K., Downie, J. A., & Poole, P. S. (2005). Role of polyhydroxybutyrate and glycogen as carbon storage compounds in pea and bean bacteroids. **Mol Plant Microbe Interact**, 18(1), 67-74. https://doi.org/10.1094/MPMI-18-0067

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. https://doi.org/10.1093/jxb/err291

Lu, P., Ma, D., Chen, Y., Guo, Y., Chen, G. Q., Deng, H., & Shi, Y. (2013). L-glutamine provides acid resistance for *Escherichia coli* through enzymatic release of ammonia. **Cell Res**, 23(5), 635-644. https://doi.org/10.1038/cr.2013.13

Lundin, E., Tang, P. C., Guy, L., Näsvall, J., & Andersson, D. I. (2018). Experimental determination and prediction of the fitness effects of random point mutations in the biosynthetic enzyme HisA. **Mol Biol Evol,** 35(3), 704-718. https://doi.org/10.1093/molbev/msx325

Lynch, M., Ackerman, M. S., Gout, J. F., Long, H., Sung, W., Thomas, W. K., & Foster, P. L. (2016). Genetic drift, selection and the evolution of the mutation rate. **Nat Rev Genet**, 17(11), 704-714. https://doi.org/10.1038/nrg.2016.104

López-Baena, F. J., Ruiz-Sainz, J. E., Rodríguez-Carvajal, M. A., & Vinardell, J. M. (2016). Bacterial molecular signals in the *Sinorhizobium fredii*-Soybean symbiosis. **Int J Mol Sci**, 17(5). https://doi.org/10.3390/ijms17050755

Macho, A. P., Guidot, A., Barberis, P., Beuzon, C. R., & Genin, S. (2010). A competitive index assay identifies several *Ralstonia solanacearum* Type III effector mutant strains with reduced fitness in host plants. **Molecular Plant-Microbe Interactions**, 23(9), 1197-1205. https://doi.org/10.1094/mpmi-23-9-1197

MacLean, A. M., Finan, T. M., & Sadowsky, M. J. (2007). Genomes of the symbiotic nitrogen-fixing bacteria of legumes. **Plant Physiology**, 144(2), 615-622. https://doi.org/10.1104/pp.107.101634

Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Stougaard, J. (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. **Nature**, 425(6958), 637-640. https://doi.org/10.1038/nature02045

Madsen, L. H., Tirichine, L., Jurkiewicz, A., Sullivan, J. T., Heckmann, A. B., Bek, A. S., Stougaard, J. (2010). The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. **Nature Communications**, 1. https://doi.org/10.1038/ncomms1009

Maillet, F., Fournier, J., Mendis, H. C., Tadege, M., Wen, J., Ratet, P., Jones, K. M. (2020). *Sinorhizobium meliloti* succinylated high-molecular-weight succinoglycan and the *Medicago truncatula* LysM receptor-like kinase MtLYK10 participate independently in symbiotic infection. **Plant J**, 102(2), 311-326. https://doi.org/10.1111/tpj.14625

Malavolta, E. (1980). Elementos de Nutrição Mineral de Plantas. In: Editora Agronômica Ceres.

Manriquez, B., Muller, D., & Prigent-Combaret, C. (2021). Experimental evolution in plant-microbe systems: a tool for deciphering the functioning and evolution of plant-associated microbial communities. **Front Microbiol**, 12, 619122. https://doi.org/10.3389/fmicb.2021.619122

Marchetti, M., Capela, D., Glew, M., Cruveiller, S., Chane-Woon-Ming, B., Gris, C., Masson-Boivin, C. (2010). Experimental evolution of a plant pathogen into a legume symbiont. **Plos Biology**, 8(1). https://doi.org/e1000280/10.1371/journal.pbio.1000280

Marchetti, M., Catrice, O., Batut, J., & Masson-Boivin, C. (2011). *Cupriavidus taiwanensis* bacteroids in *Mimosa pudica* indeterminate nodules are not terminally differentiated. **Applied and Environmental Microbiology**, 77(6), 2161-2164. https://doi.org/10.1128/aem.02358-10

Marchetti, M., Clerissi, C., Yousfi, Y., Gris, C., Bouchez, O., Rocha, E., Masson-Boivin, C. (2017). Experimental evolution of rhizobia may lead to either extra- or intracellular symbiotic adaptation depending on the selection regime. **Mol Ecol**, 26(7), 1818-1831. https://doi.org/10.1111/mec.13895

Marchetti, M., Jauneau, A., Capela, D., Remigi, P., Gris, C., Batut, J., & Masson-Boivin, C. (2014). Shaping bacterial symbiosis with legumes by experimental evolution. **Molecular Plant-Microbe Interactions**, 27(9), 956-964. https://doi.org/10.1094/mpmi-03-14-0083-r

Marenda, M., Brito, B., Callard, D., Genin, S., Barberis, P., Boucher, C., & Arlat, M. (1998). PrhA controls a novel regulatory pathway required for the specific induction of *Ralstonia solanacearum hrp* genes in the presence of plant cells. **Molecular Microbiology**, 27(2), 437-453. https://doi.org/10.1046/j.1365-2958.1998.00692.x

Marie, C., Deakin, W. J., Viprey, V., Kopciñska, J., Golinowski, W., Krishnan, H. B., Broughton, W. J. (2003). Characterization of Nops, nodulation outer proteins, secreted via the type III secretion system of NGR234. **Mol Plant Microbe Interact**, 16(9), 743-751. https://doi.org/10.1094/MPMI.2003.16.9.743

Marin, J., Battistuzzi, F. U., Brown, A. C., & Hedges, S. B. (2017). The timetree of Prokaryotes: new insights into their evolution and speciation. **Mol Biol Evol**, 34(2), 437-446. https://doi.org/10.1093/molbev/msw245

Marino, D., Andrio, E., Danchin, E. G., Oger, E., Gucciardo, S., Lambert, A., Pauly, N. (2011). A *Medicago truncatula* NADPH oxidase is involved in symbiotic nodule functioning. **New Phytol**, 189(2), 580-592. https://doi.org/10.1111/j.1469-8137.2010.03509.x

Martinez, E., Palacios, R., & Sanchez, F. (1987). Nitrogen-fixing nodules induced by *Agrobacterium tumefaciens* harboring *Rhizobium phaseoli* plasmids. **Journal of Bacteriology**, 169(6), 2828-2834.

Martínez-Hidalgo, P., & Hirsch, A. (2017). The nodule microbiome: N_2-fixing rhizobia do not live alone. **Phytobiomes**.

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

Masson-Boivin, C., & Sachs, J. L. (2018). Symbiotic nitrogen fixation by rhizobia-the roots of a success story. **Curr Opin Plant Biol**, 44, 7-15. https://doi.org/10.1016/j.pbi.2017.12.001

Mayhood, P., & Mirza, B. S. (2021). Soybean root nodule and rhizosphere microbiome: distribution of rhizobial and nonrhizobial endophytes. **Appl Environ Microbiol**, 87(10). https://doi.org/10.1128/AEM.02884-20

McDonald, M. J., Hsieh, Y. Y., Yu, Y. H., Chang, S. L., & Leu, J. Y. (2012). The evolution of low mutation rates in experimental mutator populations of *Saccharomyces cerevisiae*. **Curr Biol**, 22(13), 1235-1240. https://doi.org/10.1016/j.cub.2012.04.056

Mei, P.-P., Wang, P., Yang, H., Gui, L.-G., Christie, P., & Li, L. (2021). Maize/faba bean intercropping with rhizobial inoculation in a reclaimed desert soil enhances productivity and symbiotic N_2 fixation and reduces apparent N losses. **Soil & tillage research**, v. 213.

Mello, F. d. A. F., Brasil Sobrinho, M. d. O. C. d., & Arzolla, S. (1989). Fertilidade do solo. In (pp. 400). Nobel.

Mellor, R. B. (1989). Bacteroids in the rhizobium -legume symbiosis inhabit a plant internal lytic compartment: implications for other microbial endosymbioses. **Journal of Experimental Botany**, v. 40, pp. 831–839.

Melnyk, R. A., Hossain, S. S., & Haney, C. H. (2019). Convergent gain and loss of genomic islands drive lifestyle changes in plant-associated *Pseudomonas*. **ISME J**, 13(6), 1575-1588. https://doi.org/10.1038/s41396-019-0372-5

Menegat, S., Ledo, A., & Tirado, R. (2022). Greenhouse gas emissions from global production and use of nitrogen synthetic fertilisers in agriculture. **Sci Rep**, 12(1), 14490. https://doi.org/10.1038/s41598-022-18773-w

Meng, F. H., Yao, J., & Allen, C. (2011). A MotN mutant of *Ralstonia solanacearum* is hypermotile and has reduced virulence. **Journal of Bacteriology**, 193(10), 2477-2486. https://doi.org/10.1128/jb.01360-10

Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., & Kondorosi, E. (2003). A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. **Plant Physiology**, 132(1), 161-173. https://doi.org/10.1104/pp.102.018192

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Kondorosi, E. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. **Proc Natl Acad Sci U S A**, 103(13), 5230-5235.

Miller-Williams, M., Loewen, P. C., & Oresnik, I. J. (2006). Isolation of salt-sensitive mutants of *Sinorhizobium meliloti* strain Rm1021. **Microbiology**, 152(Pt 7), 2049-2059. https://doi.org/10.1099/mic.0.28937-0

Mishra, R. P. N., Tisseyre, P., Melkonian, R., Chaintreuil, C., Miche, L., Klonowska, A., Moulin, L. (2012). Genetic diversity of *Mimosa pudica* rhizobial symbionts in soils of French Guiana: investigating the origin and diversity of *Burkholderia phymatum* and other beta-rhizobia. **Fems Microbiology Ecology**, 79(2), 487-503. https://doi.org/10.1111/j.1574-6941.2011.01235.x

Moling, S., Pietraszewska-Bogiel, A., Postma, M., Fedorova, E., Hink, M. A., Limpens, E., Bisseling, T. (2014). Nod factor receptors form heteromeric complexes and are essential for intracellular infection in medicago nodules. **Plant Cell**, 26(10), 4188-4199. https://doi.org/10.1105/tpc.114.129502

Montiel, J., Downie, J. A., Farkas, A., Bihari, P., Herczeg, R., Bálint, B., Kondorosi, É. (2017). Morphotype of bacteroids in different legumes correlates with the number and type of

symbiotic NCR peptides. **Proc Natl Acad Sci U S A**, 114(19), 5041-5046. https://doi.org/10.1073/pnas.1704217114

Montiel, J., Szűcs, A., Boboescu, I. Z., Gherman, V. D., Kondorosi, É., & Kereszt, A. (2016). Terminal bacteroid differentiation is associated with variable morphological changes in legume species belonging to the Inverted Repeat-Lacking Clade. **Mol Plant Microbe Interact**, 29(3), 210-219. https://doi.org/10.1094/MPMI-09-15-0213-R

Moreira, F. M. S., & Siqueira, J. O. (2006). Microbiologia e bioquímica do solo. Universidade Federal de Lavras. UFLA. 744p

Mori, Y., Ishikawa, S., Ohnishi, H., Shimatani, M., Morikawa, Y., Hayashi, K., Hikichi, Y. (2017). Involvement of ralfuranones in the quorum sensing signalling pathway and virulence of *Ralstonia solanacearum* strain OE1-1. **Mol Plant Pathol**. https://doi.org/10.1111/mpp.12537

Mori, Y., Ishikawa, S., Ohnishi, H., Shimatani, M., Morikawa, Y., Hayashi, K., Hikichi, Y. (2018). Involvement of ralfuranones in the quorum sensing signalling pathway and virulence of *Ralstonia solanacearum* strain OE1-1. **Mol Plant Pathol**, 19(2), 454-463. https://doi.org/10.1111/mpp.12537

Moulin, L., Munive, A., Dreyfus, B., & Boivin-Masson, C. (2001). Nodulation of legumes by members of the beta-subclass of Proteobacteria. **Nature**, 411(6840), 948-950.

Mueller, U. G., & Sachs, J. L. (2015). Engineering microbiomes to improve plant and animal health. **Trends in Microbiology**, 23(10), 606-617. https://doi.org/10.1016/j.tim.2015.07.009

Murashige, T., & Skoog, F. V. O. P. (1962). A revised medium for rapid growth and bio assays with *Tobacco* tissue cultures. **Physiologia plantarum**, v.15.

Nakatsukasa, H., Uchiumi, T., Kucho, K., Suzuki, A., Higashi, S., & Abe, M. (2008). Transposon mediation allows a symbiotic plasmid of *Rhizobium leguminosarum* bv. *trifolii* to become a symbiosis island in *Agrobacterium* and *Rhizobium*. **J Gen Appl Microbiol**, 54(2), 107-118. https://doi.org/10.2323/jgam.54.107

Nandasena, K. G., O'Hara, G. W., Tiwari, R. P., & Howieson, J. G. (2006). Rapid *in situ* evolution of nodulating strains for *Biserrula pelecinus* L. through lateral transfer of a symbiosis island from the original mesorhizobial inoculant. **Applied and Environmental Microbiology**, 72(11), 7365-7367. https://doi.org/10.1128/aem.00889.06

Nandasena, K. G., O'Hara, G. W., Tiwari, R. P., Sezmis, E., & Howieson, J. G. (2007). *In situ* lateral transfer of symbiosis islands results in rapid evolution of diverse competitive strains of mesorhizobia suboptimal in symbiotic nitrogen fixation on the pasture legume *Biserrula pelecinus* L. **Environmental Microbiology**, 9(10), 2496-2511. https://doi.org/10.1111/j.1462-2920.2007.01368.x

Newman, J. D., Schultz, B. W., & Noel, K. D. (1992). Dissection of nodule development by supplementation of *Rhizobium leguminosarum* biovar *phaseoli* purine auxotrophs with 4-aminoimidazole-5-carboxamide riboside. **Plant Physiol**, 99(2), 401-408. https://doi.org/10.1104/pp.99.2.401

Nguyen Ba, A. N., Cvijović, I., Rojas Echenique, J. I., Lawrence, K. R., Rego-Costa, A., Liu, X., Desai, M. M. (2019). High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. **Nature**, 575(7783), 494-499. https://doi.org/10.1038/s41586-019-1749-3

Nicoud, Q., Lamouche, F., Chaumeret, A., Balliau, T., Le Bars, R., Bourge, M., Alunni, B. (2021). *Bradyrhizobium diazoefficiens* USDA110 nodulation of *Aeschynomene afraspera* is associated with atypical terminal bacteroid differentiation and suboptimal symbiotic efficiency. **mSystems,** 6(3). https://doi.org/10.1128/mSystems.01237-20

Noel, K. D., Forsberg, L. S., & Carlson, R. W. (2000). Varying the abundance of O antigen in *Rhizobium etli* and its effect on symbiosis with *Phaseolus vulgaris*. **J Bacteriol**, 182(19), 5317-5324.

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. **Sci Rep**, 7(1), 448. https://doi.org/10.1038/s41598-017-00559-0

Novikova, N., & Safronova, V. (1992). Transconjugants of *Agrobacterium radiobacter* harbouring sym genes of *Rhizobium galegae* can form an effective symbiosis with Medicago sativa. **FEMS Microbiol Lett**, 72(3), 261-268. https://doi.org/10.1016/0378-1097(92)90472-z

Ohwada, T., Sasaki, Y., Koike, H., Igawa, K., & Sato, T. (1998). Correlation between NaCl sensitivity of *Rhizobium* bacteria and ineffective nodulation of leguminous plants. **Biosci Biotechnol Biochem**, 62(11), 2086-2090. https://doi.org/10.1271/bbb.62.2086

Okazaki, S., Kaneko, T., Sato, S., & Saeki, K. (2013). Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. **Proc Natl Acad Sci U S A**, 110(42), 17131-17136. https://doi.org/10.1073/pnas.1302360110

Okazaki, S., Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., Giraud, E. (2016). *Rhizobium*-legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. **ISME J**, 10(1), 64-74. https://doi.org/10.1038/ismej.2015.103

Okazaki, S., Zehner, S., Hempel, J., Lang, K., & Göttfert, M. (2009). Genetic organization and functional analysis of the type III secretion system of *Bradyrhizobium elkanii*. **FEMS Microbiol Lett**, 295(1), 88-95. https://doi.org/10.1111/j.1574-6968.2009.01593.x

Oldroyd, G. E. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. **Nat Rev Microbiol**, 11(4), 252-263. https://doi.org/10.1038/nrmicro2990

Oldroyd, G. E. D., & Downie, J. M. (2008). Coordinating nodule morphogenesis with rhizobial infection in legumes. **Annual Review of Plant Biology**, 59, 519-546. https://doi.org/10.1146/annurev.arplant.59.032607.092839

Oldroyd, G. E. D., Murray, J. D., Poole, P. S., & Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. In B. L. Bassler, M. Lichten, & G. Schupbach (Eds.), **Annual Review Genetics**, v. 45, p. 119-144. https://doi.org/10.1146/annurev-genet-110410-132549

Oono, R., Anderson, C. G., & Denison, R. F. (2011). Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. **Proceedings of the Royal Society B-Biological Sciences**, 278(1718), 2698-2703. https://doi.org/10.1098/rspb.2010.2193

Oono, R., Denison, R. F., & Kiers, E. T. (2009). Controlling the reproductive fate of rhizobia: how universal are legume sanctions? **New Phytologist**, 183(4), 967-979. https://doi.org/10.1111/j.1469-8137.2009.02941.x

Op den Camp, R. H. M., Polone, E., Fedorova, E., Roelofsen, W., Squartini, A., Op den Camp, H. J. M., Geurts, R. (2012). Nonlegume *Parasponia andersonii* deploys a broad *Rhizobium* host range strategy resulting in Largely Variable Symbiotic Effectiveness. **Molecular Plant-Microbe Interactions**, 25(7), 954-963. https://doi.org/10.1094/mpmi-11-11-0304

Oresnik, I. J., Liu, S. L., Yost, C. K., & Hynes, M. F. (2000). Megaplasmid pRme2011a of *Sinorhizobium meliloti* is not required for viability. **J Bacteriol**, 182(12), 3582-3586. https://doi.org/10.1128/JB.182.12.3582-3586.2000

Ott, T., Sullivan, J., James, E. K., Flemetakis, E., Günther, C., Gibon, Y., Udvardi, M. (2009). Absence of symbiotic leghemoglobins alters bacteroid and plant cell differentiation during development of *Lotus japonicus* root nodules. **Mol Plant Microbe Interact**, 22(7), 800-808. https://doi.org/10.1094/MPMI-22-7-0800

Ott, T., van Dongen, J. T., Günther, C., Krusell, L., Desbrosses, G., Vigeolas, H., Udvardi, M. K. (2005). Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. **Curr Biol**, 15(6), 531-535. https://doi.org/10.1016/j.cub.2005.01.042

Pankey, S. M., Foxall, R. L., Ster, I. M., Perry, L. A., Schuster, B. M., Donner, R. A., Whistler, C. A. (2017). Host-selected mutations converging on a global regulator drive an adaptive leap towards symbiosis in bacteria. **Elife**, 6. https://doi.org/10.7554/eLife.24414

Pankhurst, C. E., & Biggs, D. R. (1980). Sensitivity of *Rhizobium* to selected isoflavonoids. **Can J Microbiol**, 26(4), 542-545. https://doi.org/10.1139/m80-092

Panter, S., Thomson, R., de Bruxelles, G., Laver, D., Trevaskis, B., & Udvardi, M. (2000). Identification with proteomics of novel proteins associated with the peribacteroid membrane of soybean root nodules. **Mol Plant Microbe Interact**, 13(3), 325-333. https://doi.org/10.1094/MPMI.2000.13.3.325

Parniske, M. (2018). Uptake of bacteria into living plant cells, the unifying and distinct feature of the nitrogen-fixing root nodule symbiosis. **Curr Opin Plant Biol**, 44, 164-174. https://doi.org/10.1016/j.pbi.2018.05.016

Peck, M. C., Fisher, R. F., & Long, S. R. (2006). Diverse flavonoids stimulate NodD1 binding to nod gene promoters in *Sinorhizobium meliloti*. **J Bacteriol**, 188(15), 5417-5427. https://doi.org/10.1128/JB.00376-06

Peeters, N., Carrere, S., Anisimova, M., Plener, L., Cazale, A.-C., & Genin, S. (2013). Repertoire, unified nomenclature and evolution of the Type III effector gene set in the *Ralstonia solanacearum* species complex. **Bmc Genomics**, 14, Article 859. https://doi.org/10.1186/1471-2164-14-859

Peeters, N., Guidot, A., Vailleau, F., & Valls, M. (2013). *Ralstonia solanacearum*, a widespread bacterial plant pathogen in the post-genomic era. **Mol Plant Pathol**, 14(7), 651-662. https://doi.org/10.1111/mpp.12038

Pelaez-Vico, M., Bernabéu-Roda, L., Kohlen, W., Soto, M., & Lopez-Raez, J. (2016). Strigolactones in the *Rhizobium*-legume symbiosis: stimulatory effect on bacterial surface motility and down-regulation of their levels in nodulated plants. **Plant Science**, v. 11, p. 119-127.

Peltzer, M. D., Roques, N., Poinsot, V., Aguilar, O. M., Batut, J., & Capela, D. (2008). Auxotrophy accounts for nodulation defect of most *Sinorhizobium meliloti* mutants in the branched-

chain amino acid biosynthesis pathway. **Molecular Plant-Microbe Interactions**, 21(9), 1232-1241. https://doi.org/10.1094/mpmi-21-9-1232

Penterman, J., Abo, R. P., De Nisco, N. J., Arnold, M. F., Longhi, R., Zanda, M., & Walker, G. C. (2014). Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. **Proc Natl Acad Sci U S A**, 111(9), 3561-3566. https://doi.org/10.1073/pnas.1400450111

Perret, X., Staehelin, C., & Broughton, W. J. (2000). Molecular basis of symbiotic promiscuity. **Microbiology and Molecular Biology Reviews**, 64(1), 180.

Perrier, A., Barlet, X., Peyraud, R., Rengel, D., Guidot, A., & Genin, S. (2018). Comparative transcriptomic studies identify specific expression patterns of virulence factors under the control of the master regulator PhcA in the *Ralstonia solanacearum* species complex. **Microb Pathog**, 116, 273-278. https://doi.org/10.1016/j.micpath.2018.01.028

Perrier, A., Peyraud, R., Rengel, D., Barlet, X., Lucasson, E., Gouzy, J., Guidot, A. (2016). Enhanced *in planta* fitness through adaptive mutations in EfpR, a dual regulator of virulence and metabolic functions in the plant pathogen *Ralstonia solanacearum*. **PLoS Pathog**, 12(12), e1006044. https://doi.org/10.1371/journal.ppat.1006044

Persson. (2015). *Candidatus Frankia Datiscae* Dg1, the actinobacterial microsymbiont of *Datisca glomerata*, expresses the canonical *nod* genes *nodABC* in symbiosis with its host plant. **PLoS One**.

Peyraud, R., Cottret, L., Marmiesse, L., Gouzy, J., & Genin, S. (2016). A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. **PLoS Pathog**, 12(10), e1005939. https://doi.org/10.1371/journal.ppat.1005939

Peyraud, R., Denny, T., & Genin, S. (2017). Exopolysaccharide quantification for the plant pathogen *Ralstonia solanacearum*. **Bio-Protocol**, v. 7, p. 1-8.

Philippe, N., Crozat, E., Lenski, R. E., & Schneider, D. (2007). Evolution of global regulatory networks during a long-term experiment with *Eisherichia coli*. **Bioessays**, 29(9), 846-860. https://doi.org/10.1002/bies.20629

Pierre, O., Engler, G., Hopkins, J., Brau, F., Boncompagni, E., & Hérouart, D. (2013). Peribacteroid space acidification: a marker of mature bacteroid functioning in *Medicago truncatula* nodules. **Plant Cell Environ**, 36(11), 2059-2070. https://doi.org/10.1111/pce.12116

Plener, L., Manfredi, P., Valls, M., & Genin, S. (2010). PrhG, a Transcriptional regulator responding to growth conditions, is involved in the control of the Type III Secretion System regulon in *Ralstonia solanacearum*. **Journal of Bacteriology**, 192(4), 1011-1019. https://doi.org/10.1128/jb.01189-09

Pobigaylo, N., Szymczak, S., Nattkemper, T. W., & Becker, A. (2008). Identification of genes relevant to symbiosis and competitiveness in *Sinorhizobium meliloti* using signature-tagged mutants. **Mol Plant Microbe Interact**, 21(2), 219-231. https://doi.org/10.1094/MPMI-21-2-0219

Poole, P., Ramachandran, V., & Terpolilli, J. (2018). Rhizobia: from saprophytes to endosymbionts. **Nat Rev Microbiol**, 16(5), 291-303. https://doi.org/10.1038/nrmicro.2017.171

Poueymiro, M., Cunnac, S., Barberis, P., Deslandes, L., Peeters, N., Cazale-Noel, A. C., Genin, S. (2009). Two Type III Secretion System effectors from *Ralstonia solanacearum* GMI1000 determine host-range specificity on *Tobacco*. **Molecular Plant-Microbe Interactions**, 22(5), 538-550. https://doi.org/10.1094/mpmi-22-5-0538

Poueymiro, M., & Genin, S. (2009). Secreted proteins from *Ralstonia solanacearum*: a hundred tricks to kill a plant. **Current Opinion in Microbiology**, 12(1), 44-52. https://doi.org/10.1016/j.mib.2008.11.008

Preisig, O., Zufferey, R., Thöny-Meyer, L., Appleby, C. A., & Hennecke, H. (1996). A high-affinity *cbb3*-type cytochrome oxidase terminates the symbiosis-specific respiratory chain of *Bradyrhizobium japonicum*. **J Bacteriol**, 178(6), 1532-1538. https://doi.org/10.1128/jb.178.6.1532-1538.1996

Prell, J., & Poole, P. (2006). Metabolic changes of rhizobia in legume nodules. **Trends in Microbiology**, 14(4), 161-168. https://doi.org/10.1016/j.tim.2006.02.005

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. **Proc Natl Acad Sci U S A**, 106(30), 12477-12482. https://doi.org/10.1073/pnas.0903653106

Prell, J., Whiteb, J. P., Bourdesa, A., Bunnewella, S., Bongaertsc, R. J., & Poole, P. (2009). Legumes regulate *Rhizobium* bacteroid development and persistence by the supply of branched-chain amino acids. **Proceedings of the National Academy of Sciences**, v. 106, p. 12477–12482.

Price, P. A., Tanner, H. R., Dillon, B. A., Shabab, M., Walker, G. C., & Griffitts, J. S. (2015). Rhizobial peptidase HrrP cleaves host-encoded signaling peptides and mediates symbiotic compatibility. **Proceedings of the National Academy of Sciences of the United States of America**, 112(49), 15244-15249. https://doi.org/10.1073/pnas.1417797112

Pueppke, S. G., Bolanos-Vasquez, M. C., Werner, D., Bec-Ferte, M. P., Prome, J. C., & Krishnan, H. B. (1998). Release of flavonoids by the soybean cultivars McCall and peking and their perception as signals by the nitrogen-fixing symbiont *Sinorhizobium fredii*. **Plant Physiol**, 117(2), 599-606. https://doi.org/10.1104/pp.117.2.599

Pueppke, S. G., & Broughton, W. J. (1999). *Rhizobium* sp. strain NGR234 and *R-fredii* USDA257 share exceptionally broad, nested host ranges. **Molecular Plant-Microbe Interactions**, 12(4), 293-318. https://doi.org/10.1094/mpmi.1999.12.4.293

Pérez Carrascal, O. M., VanInsberghe, D., Juárez, S., Polz, M. F., Vinuesa, P., & González, V. (2016). Population genomics of the symbiotic plasmids of sympatric nitrogen-fixing *Rhizobium* species associated with *Phaseolus vulgaris*. **Environ Microbiol**, 18(8), 2660-2676. https://doi.org/10.1111/1462-2920.13415

Quides, K. W., Weisberg, A. J., Trinh, J., Salaheldine, F., Cardenas, P., Lee, H. H., Sachs, J. L. (2021). Experimental evolution can enhance benefits of rhizobia to novel legume hosts. **Proc Biol Sci**, 288(1951), 20210812. https://doi.org/10.1098/rspb.2021.0812

Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Grønlund, M., 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

Raich, J. W., Russell, A. E., Kitayama, K., Parton, W. J., & Vitousek, P. M. (2006). Temperature influences carbon accumulation in moist tropical forests. **Ecology**, 87(1), 76-87. https://doi.org/10.1890/05-0023

Ratcliff, W. C., Kadam, S. V., & Denison, R. F. (2008). Poly-3-hydroxybutyrate (PHB) supports survival and reproduction in starving rhizobia. **FEMS Microbiol Ecol**, 65(3), 391-399. https://doi.org/10.1111/j.1574-6941.2008.00544.x

Rathbun, E. A., Naldrett, M. J., & Brewin, N. J. (2002). Identification of a family of extensin-like glycoproteins in the lumen of rhizobium-induced infection threads in pea root nodules. **Mol Plant Microbe Interact**, 15(4), 350-359. https://doi.org/10.1094/MPMI.2002.15.4.350

Raynes, Y., & Sniegowski, P. D. (2014). Experimental evolution and the dynamics of genomic mutation rate modifiers. **Heredity** (Edinb), 113(5), 375-380. https://doi.org/10.1038/hdy.2014.49

Regus, J. U., Gano, K. A., Hollowell, A. C., & Sachs, J. L. (2014). Efficiency of partner choice and sanctions in *Lotus* is not altered by nitrogen fertilization. **Proceedings of the Royal Society B-Biological Sciences**, 281(1781), Article 20132587. https://doi.org/10.1098/rspb.2013.2587

Regus, J. U., Quides, K. W., O'Neill, M. R., Suzuki, R., Savory, E. A., Chang, J. H., & Sachs, J. L. (2017). Cell autonomous sanctions in legumes target ineffective rhizobia in nodules with mixed infections. **Am J Bot**, 104(9), 1299-1312. https://doi.org/10.3732/ajb.1700165

Remigi. (2016). Symbiosis within Symbiosis: Evolving Nitrogen-Fixing Legume Symbionts. **Trends in Microbiology.**

Remigi, P., Anisimova, M., Guidot, A., Genin, S., & Peeters, N. (2011). Functional diversification of the GALA type III effector family contributes to *Ralstonia solanacearum* adaptation on different plant hosts. **New Phytologist**, 192(4), 976-987. https://doi.org/10.1111/j.1469-8137.2011.03854.x

Remigi, P., Capela, D., Clerissi, C., Tasse, L., Torchet, R., Bouchez, O., Masson-Boivin, C. (2014). Transient hypermutagenesis accelerates the evolution of legume endosymbionts following horizontal gene transfer. **PLoS Biol**, 12(9), e1001942. https://doi.org/10.1371/journal.pbio.1001942

Reynolds, P. H., Boland, M. J., Blevins, D. G., Schubert, K. R., & Randall, D. D. (1982). Enzymes of amide and ureide biogenesis in developing soybean nodules. **Plant Physiol**, 69(6), 1334-1338. https://doi.org/10.1104/pp.69.6.1334

Robinson, C. D., Klein, H. S., Murphy, K. D., Parthasarathy, R., Guillemin, K., & Bohannan, B. J. M. (2018). Experimental bacterial adaptation to the zebrafish gut reveals a primary role for immigration. **PLoS Biol**, 16(12), e2006893. https://doi.org/10.1371/journal.pbio.2006893

Rogel, M., Hernández-Lucas, I., Kuykendall, L., Balkwill, D., & Martinez-Romero, E. (2001). Nitrogen-fixing nodules with *Ensifer adhaerens* harboring *Rhizobium tropici* symbiotic plasmids. **Appl Environ Microbiol,** 67(7), 3264-3268.

Rogel, M. A., Ormeno-Orrillo, E., & Romero, E. M. (2011). Symbiovars in rhizobia reflect bacterial adaptation to legumes. **Systematic and Applied Microbiology**, 34(2), 96-104. https://doi.org/10.1016/j.syapm.2010.11.015

Ronson, C., Lyttleton, P., & Robertson, J. (1981). C4-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. In (Vol. 78, pp. 4284-4288). **Proceedings of the National Academy of Sciences**.

Ronson, C. W., Astwood, P. M., & Downie, J. A. (1984). Molecular cloning and genetic organization of C4-dicarboxylate transport genes from *Rhizobium leguminosarum*. **J Bacteriol**, 160(3), 903-909. https://doi.org/10.1128/jb.160.3.903-909.1984

Ronson, C. W., Lyttleton, P., & Robertson, J. G. (1981). C(4)-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. **Proc Natl Acad Sci U S A**, 78(7), 4284-4288. https://doi.org/10.1073/pnas.78.7.4284

Roswall, T. (1983). Biogeochemical Cycles and Their Interactions. In (pp. 46-50): John Wiley & Sons: Chichester,.

Roth, L. E., & Stacey, G. (1989). Bacterium release into host cells of nitrogen-fixing soybean nodules: the symbiosome membrane comes from three sources. **Eur J Cell Biol**, 49(1), 13-23.

Roux, B., Rodde, N., Jardinaud, M. F., Timmers, T., Sauviac, L., Cottret, L., Gamas, P. (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. **Plant J**, 77(6), 817-837. https://doi.org/10.1111/tpj.12442

Rutten, P. J., & Poole, P. S. (2019). Oxygen regulatory mechanisms of nitrogen fixation in rhizobia. Adv Microb Physiol, 75, 325-389. https://doi.org/10.1016/bs.ampbs.2019.08.001

Saad, M. M., Crevecoeur, M., Masson-Boivin, C., & Perret, X. (2012). The Type 3 Protein Secretion System of *Cupriavidus taiwanensis* strain LMG19424 compromises symbiosis with *Leucaena leucocephala*. **Applied and Environmental Microbiology**, 78(20), 7476-7479. https://doi.org/10.1128/aem.01691-12

Sabbioni, G., & Forlani, G. (2022). The emerging role of proline in the establishment and functioning of legume-. **Front Plant Sci**, 13, 888769. https://doi.org/10.3389/fpls.2022.888769

Sachs, J. L., Skophammer, R. G., Bansal, N., & Stajich, J. E. (2014). Evolutionary origins and diversification of proteobacterial mutualists. **Proceedings of the Royal Society B-Biological Sciences**, 281(1775), 9. Article 20132146. https://doi.org/10.1098/rspb.2013.2146

Saeki, K. (2011). Rhizobial measures to evade host defense strategies and endogenous threats to persistent symbiotic nitrogen fixation: a focus on two legume-rhizobium model systems. **Cellular and Molecular Life Sciences**, 68(8), 1327-1339. https://doi.org/10.1007/s00018-011-0650-5

Sahlman, K. (1963). An electron microscope study of root hair infection by Rhizobium. **J Gen Microbiol**, 33, 425-427. https://doi.org/10.1099/00221287-33-3-425

Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., Boucher, C. A. (2002). Genome sequence of the plant pathogen *Ralstonia solanacearum*. **Nature**, 415(6871), 497-502.

Salas, M. E., Lozano, M. J., López, J. L., Draghi, W. O., Serrania, J., Torres Tejerizo, G. A., Lagares, A. (2017). Specificity traits consistent with legume-rhizobia coevolution displayed by

Ensifer meliloti rhizosphere colonization. **Environ Microbiol**, 19(9), 3423-3438. https://doi.org/10.1111/1462-2920.13820

Salminen, S., & Streeter, J. (1990). Factors contributing to the accumulation of glutamate in *Bradyrhizobium japonicum* bacteroids under microaerobic conditions. **Journal of General Microbiology**, v. 136.

Sankari, S., Babu, V. M. P., Bian, K., Alhhazmi, A., Andorfer, M. C., Avalos, D. M., Walker, G. C. (2022). A haem-sequestering plant peptide promotes iron uptake in symbiotic bacteria. **Nat Microbiol**, 7(9), 1453-1465. https://doi.org/10.1038/s41564-022-01192-y

Schell, M. A. (2000). Control of virulence and pathogenicity genes of *Ralstonia solanacearum* by an elaborate sensory network. **Annual Review of Phytopathology**, 38, 263-292.

Schiessl, K., Lilley, J. L. S., Lee, T., Tamvakis, I., Kohlen, W., Bailey, P. C., Oldroyd, G. E. D. (2019). Nodule Inception recruits the lateral root developmental program for symbiotic nodule organogenesis in *Medicago truncatula*. **Curr Biol**, 29(21), 3657-3668.e3655. https://doi.org/10.1016/j.cub.2019.09.005

Schlaman, H. R., Okker, R. J., & Lugtenberg, B. J. (1992). Regulation of nodulation gene expression by NodD in rhizobia. **J Bacteriol**, 174(16), 5177-5182.

Schulte, C. C. M., Borah, K., Wheatley, R. M., Terpolilli, J. J., Saalbach, G., Crang, N., Poole, P. S. (2021). Metabolic control of nitrogen fixation in rhizobium-legume symbioses. **Sci Adv**, 7(31). https://doi.org/10.1126/sciadv.abh2433

Schulte, C. C. M., Ramachandran, V. K., Papachristodoulou, A., & Poole, P. S. (2022). Genome-scale metabolic modelling of lifestyle changes in *Rhizobium leguminosarum*. **mSystems**, 7(1), e0097521. https://doi.org/10.1128/msystems.00975-21

Schuster, B. M., Perry, L. A., Vaughn S. Cooper, V. S., & Whistler, C. A. (2010). Breaking the language barrier: experimental evolution of non-native *Vibrio fischeri* in squid tailors luminescence to the host. **Symbiosis**, v. 51.

Schwedock, J. S., & Long, S. R. (1992). *Rhizobium meliloti* genes involved in sulfate activation – the 2 copies of *nodPQ* and a new locus, SAA. **Genetics,** 132(4), 899-909.

Sciotti, M. A., Chanfon, A., Hennecke, H., & Fischer, H. M. (2003). Disparate oxygen responsiveness of two regulatory cascades that control expression of symbiotic genes in *Bradyrhizobium japonicum*. **J Bacteriol**, 185(18), 5639-5642.

Sena, P., Nascimento, T., Lino, J., Oliveira, G., Ferreira Neto, R., Freitas, A., Martins, L. (2020). Molecular, physiological, and symbiotic characterization of cowpea *Rhizobia* from soils under different agricultural systems in the semiarid region of Brazil. **Journal of Soil Science and Plant Nutrition**.

Shiraishi, A., Matsushita, N., & Hougetsu, T. (2010). Nodulation in black locust by the Gammaproteobacteria *Pseudomonas* sp and the Betaproteobacteria *Burkholderia* sp. **Systematic and Applied Microbiology**, 33(5), 269-274. https://doi.org/10.1016/j.syapm.2010.04.005

Simms, E. L., & Taylor, D. L. (2002). Partner choice in nitrogen-fixation mutualisms of legumes and rhizobia. **Integr Comp Biol**, 42(2), 369-380. https://doi.org/10.1093/icb/42.2.369

Skagia, A., Zografou, C., Vezyri, E., Venieraki, A., Katinakis, P., & Dimou, M. (2016). Cyclophilin PpiB is involved in motility and biofilm formation via its functional association with certain proteins. **Genes Cells**, 21(8), 833-851. https://doi.org/10.1111/gtc.12383

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 Physiol**, 145(1), 183-191. https://doi.org/10.1104/pp.107.100495

Soloveichik, G. (2019). Electrochemical synthesis of ammonia as a potential alternative to the Haber–Bosch process. In (Vol. 2). Nat Catal.

Soupene, E., Foussard, M., Boistard, P., Truchet, G., & Batut, J. (1995). Oxygen as a key developmental regulator of *Rhizobium meliloti* N_2 fixation gene expression within the alfafa root nodule. **Proceedings of the National Academy of Sciences of the United States of America**, 92(9), 3759-3763.

Soyano, T., Shimoda, Y., Kawaguchi, M., & Hayashi, M. (2019). A shared gene drives lateral root development and root nodule symbiosis pathways. **Science**, 366(6468), 1021-1023. https://doi.org/10.1126/science.aax2153

Sprent, J. I., & James, E. K. (2007). Legume evolution: where do nodules and mycorrhizas fit in? **Plant Physiology**, 144(2), 575-581. https://doi.org/10.1104/pp.107.096156

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. **Mol Phylogenet Evol**, 48(3), 1131-1144. https://doi.org/10.1016/j.ympev.2008.04.032

Stevens, J. B., de Luca, N. G., Beringer, J. E., Ringer, J. P., Yeoman, K. H., & Johnston, A. W. (2000). The purMN genes of *Rhizobium leguminosarum* and a superficial link with siderophore production. **Mol Plant Microbe Interact**, 13(2), 228-231. https://doi.org/10.1094/MPMI.2000.13.2.228

Sugawara, M., Takahashi, S., Umehara, Y., Iwano, H., Tsurumaru, H., Odake, H., Minamisawa, K. (2018). Variation in bradyrhizobial NopP effector determines symbiotic incompatibility with *Rj2*-soybeans *via* effector-triggered immunity. **Nat Commun**, 9(1), 3139. https://doi.org/10.1038/s41467-018-05663-x

Sullivan, J., & Ronson, C. (1998). Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. **Proc Natl Acad Sci U S A**, 95(9), 5145-5149.

Sullivan, J. T., Patrick, H. N., Lowther, W. L., Scott, D. B., & Ronson, C. W. (1995). Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. **Proceedings of the National Academy of Sciences of the United States of America**, 92(19), 8985-8989.

Sullivan, J. T., Trzebiatowski, J. R., Cruickshank, R. W., Gouzy, J., Brown, S. D., Elliot, R. M., Ronson, C. W. (2002). Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. **Journal of Bacteriology**, 184(11), 3086-3095. https://doi.org/10.1128/jb.184.11.3086-3095.2002

Suominen, L., Roos, C., Lortet, G., Paulin, L., & Lindström, K. (2001). Identification and structure of the *Rhizobium galegae* common nodulation genes: evidence for horizontal gene transfer. **Mol Biol Evol**, 18(6), 907-916. https://doi.org/10.1093/oxfordjournals.molbev.a003891

Swings, T., Van den Bergh, B., Wuyts, S., Oeyen, E., Voordeckers, K., Verstrepen, K. J., Michiels, J. (2017). Adaptive tuning of mutation rates allows fast response to lethal stress in. **Elife**, 6. https://doi.org/10.7554/eLife.22939

Sánchez, C., Iannino, F., Deakin, W. J., Ugalde, R. A., & Lepek, V. C. (2009). Characterization of the *Mesorhizobium loti* MAFF303099 type-three protein secretion system. **Mol Plant Microbe Interact**, 22(5), 519-528. https://doi.org/10.1094/MPMI-22-5-0519

Taiz , L., & Zeiger, E. (2014). Plant Physiology and Development. In: Sinauer Associates.

Takemura, C., Senuma, W., Hayashi, K., Minami, A., Terazawa, Y., Kaneoka, C., Hikichi, Y. (2021). PhcQ mainly contributes to the regulation of quorum sensing-dependent genes, in which PhcR is partially involved, in *Ralstonia pseudosolanacearum* strain OE1-1. **Mol Plant Pathol**, 22(12), 1538-1552. https://doi.org/10.1111/mpp.13124

Tambalo, D., Yost, C., & Hynes, M. (2015). Motility and chemotaxis in the rhizobia. **Biological Nitrogen Fixation: John Wiley & Sons, Inc**., p. 337-348.

Tampakaki, A. P., Skandalis, N., Gazi, A. D., Bastaki, M. N., Sarris, P. F., Charova, S. N., Panopoulos, N. J. (2010). Playing the "Harp": evolution of our understanding of *hrp*/*hrc* genes. **Annu Rev Phytopathol**, 48, 347-370. https://doi.org/10.1146/annurev-phyto-073009-114407

Tang, M., Bouchez, O., Cruveiller, S., Masson-Boivin, C., & Capela, D. (2020). Modulation of quorum sensing as an adaptation to nodule cell infection during experimental evolution of legume symbionts. **mBio**, 11(1). https://doi.org/10.1128/mBio.03129-19

Tang, M., & Capela, D. (2020). *Rhizobium* diversity in thelight of evolution. **Advances in Botanical Research**: K:Elsevier Ltd.

Tano, J., Ripa, M. B., Tondo, M. L., Carrau, A., Petrocelli, S., Rodriguez, M. V., Orellano, E. G. (2021). Light modulates important physiological features of ***Ralstonia pseudosolanacearum*** during the colonization of tomato plants. **Nature**, v.11.

Tatsukami, Y., Nambu, M., Morisaka, H., Kuroda, K., & Ueda, M. (2013). Disclosure of the differences of *Mesorhizobium loti* under the free-living and symbiotic conditions by comparative proteome analysis without bacteroid isolation. **BMC Microbiol**, 13, 180. https://doi.org/10.1186/1471-2180-13-180

Tenaillon, O., Rodriguez-Verdugo, A., Gaut, R. L., McDonald, P., Bennett, A. F., Long, A. D., & Gaut, B. S. (2012). The molecular diversity of adaptive convergence. **Science,** 335(6067), 457-461. https://doi.org/10.1126/science.1212986

Teulet, A., Busset, N., Fardoux, J., Gully, D., Chaintreuil, C., Cartieaux, F., Giraud, E. (2019). The rhizobial type III effector ErnA confers the ability to form nodules in legumes. **Proc Natl Acad Sci U S A,** 116(43), 21758-21768. https://doi.org/10.1073/pnas.1904456116

Tian, C. F., Garnerone, A. M., Mathieu-Demaziere, C., Masson-Boivin, C., & Batut, J. (2012). Plant-activated bacterial receptor adenylate cyclases modulate epidermal infection in the *Sinorhizobium meliloti-Medicago* symbiosis. **Proceedings of the National Academy of Sciences of the United States of America**, 109(17), 6751-6756. https://doi.org/10.1073/pnas.1120260109

Tian, C. F., Zhou, Y. J., Zhang, Y. M., Li, Q. Q., Zhang, Y. Z., Li, D. F., Chen, W. X. (2012). Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. **Proceedings of the National Academy of Sciences of the United States of America**, 109(22), 8629-8634. https://doi.org/10.1073/pnas.1120436109

Tian, L., Liu, L., Xu, S., Deng, R., Wu, P., Jiang, H., Chen, Y. (2022). A d-pinitol transporter, LjPLT11, regulates plant growth and nodule development in Lotus japonicus. **J Exp Bot**, 73(1), 351-365. https://doi.org/10.1093/jxb/erab402

Tighilt, L., Boulila, F., De Sousa, B. F. S., Giraud, E., Ruiz-Argüeso, T., Palacios, J. M., Rey, L. (2021). The *Bradyrhizobium* Sp. LmicA16 Type VI Secretion System is required for efficient nodulation of *Lupinus Spp*. **Microb Ecol**. https://doi.org/10.1007/s00248-021-01892-8

Timmers, A. C. J., Soupene, E., Auriac, M. C., de Billy, F., Vasse, J., Boistard, P., & Truchet, G. (2000). Saprophytic intracellular rhizobia in alfalfa nodules. **Molecular Plant-Microbe Interactions**, 13(11), 1204-1213. https://doi.org/10.1094/mpmi.2000.13.11.1204

Torres, M. J., Hidalgo-García, A., Bedmar, E. J., & Delgado, M. J. (2013). Functional analysis of the copy 1 of the *fixNOQP* operon of *Ensifer meliloti* under free-living micro-oxic and symbiotic conditions. J **Appl Microbiol**, 114(6), 1772-1781. https://doi.org/10.1111/jam.12168

Tso, G. H. W., Reales-Calderon, J. A., Tan, A. S. M., Sem, X., Le, G. T. T., Tan, T. G., Pavelka, N. (2018). Experimental evolution of a fungal pathogen into a gut symbiont. **Science**, 362(6414), 589-595. https://doi.org/10.1126/science.aat0537

Turner, G. L., & Gibson, A. H. (1980). Measurement of nitrogen fixation by indirect means. **Methods for evaluating biological nitrogen fixation**. pp. 111-138

Turner, S. L., & Young, J. P. W. (2000). The glutamine synthetases of rhizobia: Phylogenetics and evolutionary implications. **Molecular Biology and Evolution**, 17(2), 309-319.

Uchiumi, T., Ohwada, T., Itakura, M., Mitsui, H., Nukui, N., Dawadi, P., Minamisawa, K. (2004). Expression islands clustered on the symbiosis island of the *Mesorhizobium loti* genome. **J Bacteriol**, 186(8), 2439-2448.

Udvardi, M., & Poole, P. S. (2013). Transport and metabolism in legume-Rhizobia symbioses. In S. S. Merchant (Ed.), **Annual Review of Plant Biology**, v. 64, pp. 781-805. **Annual Reviews**. https://doi.org/10.1146/annurev-arplant-050312-120235

Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Mergaert, P. (2010). Plant peptides govern terminal differentiation of Bacteria in symbiosis. **Science**, 327(5969), 1122-1126. https://doi.org/10.1126/science.1184057

Van den Bergh, B., Swings, T., Fauvart, M., & Michiels, J. (2018). Experimental design, population dynamics, and diversity in microbial experimental evolution. **Microbiol Mol Biol Rev**, 82(3). https://doi.org/10.1128/MMBR.00008-18

van Slooten, J. C., Bhuvanasvari, T. V., Bardin, S., & Stanley, J. (1992). Two C4-dicarboxylate transport systems in *Rhizobium* sp. NGR234: rhizobial dicarboxylate transport is essential for nitrogen fixation in tropical legume symbioses. **Mol Plant Microbe Interact**, 5(2), 179-186. https://doi.org/10.1094/mpmi-5-179

van Workum, W. A. T., van Slageren, S., van Brussel, A. A. N., & Kijne, J. W. (1998). Role of exopolysaccharides of *Rhizobium leguminosarum* bv. *viciae* as host plant-specific molecules required for infection thread formation during nodulation of *Vicia sativa*. In (Vol. 11, pp. 1233–1241): **Molecular Plant Microbe Interaction**.

Vandeputte, O. M., Kiendrebeogo, M., Rajaonson, S., Diallo, B., Mol, A., El Jaziri, M., & Baucher, M. (2010). Identification of catechin as one of the flavonoids from *Combretum albiflorum* bark extract that reduces the production of quorum-sensing-controlled

virulence factors in *Pseudomonas aeruginosa* PAO1. **Appl Environ Microbiol**, 76(1), 243-253. https://doi.org/10.1128/AEM.01059-09

Vasse, J., Debilly, F., Camut, S., & Truchet, G. (1990). Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfafa nodules. **Journal of Bacteriology**, 172(8), 4295-4306.

Vasseur, P., Vallet-Gely, I., Soscia, C., Genin, S., & Filloux, A. (2005). The *pel* genes of the *Pseudomonas aeruginosa* PAK strain are involved at early and late stages of biofilm formation. **Microbiology**, 151(Pt 3), 985-997. https://doi.org/10.1099/mic.0.27410-0

Vercruysse, M., Fauvart, M., Beullens, S., Braeken, K., Cloots, L., Engelen, K., Michiels, J. (2011). A comparative transcriptome analysis of *Rhizobium etli* bacteroids: specific gene expression during symbiotic nongrowth. **Mol Plant Microbe Interact**, 24(12), 1553-1561. https://doi.org/10.1094/MPMI-05-11-0140

Verma, S. C., Chowdhury, S. P., & Tripathi, A. K. (2004). Phylogeny based on 16S rDNA and *nifH* sequences of *Ralstonia taiwanensis* strains isolated from nitrogen-fixing nodules of *Mimosa pudica*, **India. Can J Microbiol**, 50(5), 313-322. https://doi.org/10.1139/w04-020

Viero, F., Bayer, C., Fontoura, S. M. V., & Moraes, R. P. d. (2014). Ammonia volatilization from nitrogen fertilizers in no-till wheat and maize in southern Brazil. **Rev. Bras. Ciênc. Solo**, v.38.

Viprey, V., Del Greco, A., Golinowski, W., Broughton, W. J., & Perret, X. (1998). Symbiotic implications of type III protein secretion machinery in Rhizobium. Mol Microbiol, 28(6), 1381-1389.

Walker, S. A., & Downie, J. A. (2000). Entry of *Rhizobium leguminosarum* bv.*viciae* into root hairs requires minimal Nod factor specificity, but subsequent infection thread growth requires *nodO* or *nodE*. **Mol Plant Microbe Interact**, 13(7), 754-762. https://doi.org/10.1094/MPMI.2000.13.7.754

Wardell, G. E., Hynes, M. F., Young, P. J., & Harrison, E. (2022). Why are rhizobial symbiosis genes mobile? **Philos Trans R Soc Lond B Biol Sci**, 377(1842), 20200471. https://doi.org/10.1098/rstb.2020.0471

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 N_2-fixation in angiosperms. **Nature Communications**, 5, Article 4087. https://doi.org/10.1038/ncomms5087

Westhoek, A., Clark, L. J., Culbert, M., Dalchau, N., Griffiths, M., Jorrin, B., Turnbull, L. A. (2021). Conditional sanctioning in a legume-. **Proc Natl Acad Sci U S A,** 118(19). 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. **Sci Rep**, 7(1), 1419. https://doi.org/10.1038/s41598-017-01634-2

Wheatley, R. M., Ford, B. L., Li, L., Aroney, S. T. N., Knights, H. E., Ledermann, R., Poole, P. S. (2020). Lifestyle adaptations of *Rhizobium* from rhizosphere to symbiosis. **Proc Natl Acad Sci U S A**, 117(38), 23823-23834. https://doi.org/10.1073/pnas.2009094117

Whitehead, L. F., & Day, D. A. (1997). The peribacteroid membrane. In (Vol. 100, pp. 30-44): **Physiologia Plantarum**.

Wielbo, J., Marek-Kozaczuk, M., Kubik-Komar, A., & Skorupska, A. (2007). Increased metabolic potential of *Rhizobium* spp. is associated with bacterial competitiveness. **Canadian Journal of Microbiology**, 53(8), 957-967. https://doi.org/10.1139/w07-053

Wielgoss, S., Barrick, J. E., Tenaillon, O., Wiser, M. J., Dittmar, W. J., Cruveiller, S., Schneider, D. (2013). Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. **Proceedings of the National Academy of Sciences of the United States of America,** 110(1), 222-227. https://doi.org/10.1073/pnas.1219574110

Wozniak, R. A. F., & Waldor, M. K. (2010). Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. **Nature Reviews Microbiology**, 8(8), 552-563. https://doi.org/10.1038/nrmicro2382

Wulandari, D., Tittabutr, P., Songwattana, P., Piromyou, P., Teamtisong, K., Boonkerd, N., Teaumroong, N. (2022). Symbiosis contribution of non-nodulating *Bradyrhizobium cosmicum* S23321 after transferal of the symbiotic plasmid pDOA9. **Microbes Environ**, 37(2). https://doi.org/10.1264/jsme2.ME22008

Xiao, T. T., Schilderink, S., Moling, S., Deinum, E. E., Kondorosi, E., Franssen, H., Bisseling, T. (2014). Fate map of *Medicago truncatula* root nodules. **Development**, 141(18), 3517-3528. https://doi.org/10.1242/dev.110775

Xie, F., Murray, J. D., Kim, J., Heckmann, A. B., Edwards, A., Oldroyd, G. E., & Downie, J. A. (2012). Legume pectate lyase required for root infection by rhizobia. **Proc Natl Acad Sci U S A**, 109(2), 633-638. https://doi.org/10.1073/pnas.1113992109

Xiong, L., Teng, J. L., Watt, R. M., Kan, B., Lau, S. K., & Woo, P. C. (2014). Arginine deiminase pathway is far more important than urease for acid resistance and intracellular survival in *Laribacter hongkongensis*: a possible result of arc gene cassette duplication. **BMC Microbiol**, 14, 42. https://doi.org/10.1186/1471-2180-14-42

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. **J Mol Evol**, 39(3), 306-314. https://doi.org/10.1007/BF00160154

Younginger, B. S., & Friesen, M. L. (2019). Connecting signals and benefits through partner choice in plant-microbe interactions. **FEMS Microbiol Lett**, 366(18). https://doi.org/10.1093/femsle/fnz217

Yurgel, S. N., & Kahn, M. L. (2005). *Sinorhizobium meliloti dctA* mutants with partial ability to transport dicarboxylic acids. **J Bacteriol**, 187(3), 1161-1172. https://doi.org/10.1128/JB.187.3.1161-1172.2005

Zdor, R., & Pueppke, S. (1990). Nodulation competitiveness of TnS-induced mutants of *Rhizobium fredii* USDA208 that are altered in motility and extracellular polysaccharide production. **Canadian Journal of Microbiology**, pp. 52-58.

Zgadzaj. (2015). A legume genetic framework controls infection of nodules by symbiotic and endophytic Bacteria. **Plos Genet**, 11(6). Doi: 10.1371/journal.pgen.1005280

Zheng, L., & Dean, D. R. (1994). Catalytic formation of a nitrogenase iron-sulfur cluster. **J Biol Chem**, 269(29), 18723-18726.

Zhukov, V., Radutoiu, S., Madsen, L. H., Rychagova, T., Ovchinnikova, E., Borisov, A., Stougaard, J. (2008). The pea Sym37 receptor kinase gene controls infection-thread initiation and nodule development. **Mol Plant Microbe Interact**, 21(12), 1600-1608. https://doi.org/10.1094/MPMI-21-12-1600

Zufferey, R., Preisig, O., Hennecke, H., & Thöny-Meyer, L. (1996). Assembly and function of the cytochrome *cbb3* oxidase subunits in *Bradyrhizobium japonicum***. J Biol Chem**, 271(15), 9114-9119. https://doi.org/10.1074/jbc.271.15.9114

Österman, J., Marsh, J., Laine, P. K., Zeng, Z., Alatalo, E., Sullivan, J. T., Lindström, K. (2014). Genome sequencing of two *Neorhizobium galegae* strains reveals a *noeT* gene responsible for the unusual acetylation of the nodulation factors. **BMC Genomics**, 15, 500. https://doi.org/10.1186/1471-2164-15-500.

Appendix

genes

MDPI

Review

Experimental Evolution of Legume Symbionts: What Have We Learnt?

Ginaini Grazielli Doin de Moura, Philippe Remigi, Catherine Masson-Boivin and Delphine Capela *

LIPM, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan 31320, France; ginaini-grazielli.doin-de-moura@inrae.fr (G.G.D.d.M.); philippe.remigi@inrae.fr (P.R.); catherine.masson@inrae.fr (C.M.-B.)

* Correspondence: delphine.capela@inrae.fr; Tel.: +33-561-285-454

Received: 3 March 2020; Accepted: 20 March 2020; Published: 23 March 2020

Abstract: Rhizobia, the nitrogen-fixing symbionts of legumes, are polyphyletic bacteria distributed in many alpha- and beta-proteobacterial genera. They likely emerged and diversified through independent horizontal transfers of key symbiotic genes. To replay the evolution of a new rhizobium genus under laboratory conditions, the symbiotic plasmid of *Cupriavidus taiwanensis* was introduced in the plant pathogen *Ralstonia solanacearum*, and the generated proto-rhizobium was submitted to repeated inoculations to the *C. taiwanensis* host, *Mimosa pudica* L. This experiment validated a two-step evolutionary scenario of key symbiotic gene acquisition followed by genome remodeling under plant selection. Nodulation and nodule cell infection were obtained and optimized mainly via the rewiring of regulatory circuits of the recipient bacterium. Symbiotic adaptation was shown to be accelerated by the activity of a mutagenesis cassette conserved in most rhizobia. Investigating mutated genes led us to identify new components of *R. solanacearum* virulence and *C. taiwanensis* symbiosis. Nitrogen fixation was not acquired in our short experiment. However, we showed that post-infection sanctions allowed the increase in frequency of nitrogen-fixing variants among a non-fixing population in the *M. pudica*–*C. taiwanensis* system and likely allowed the spread of this trait in natura. Experimental evolution thus provided new insights into rhizobium biology and evolution.

Keywords: rhizobia; experimental evolution; nitrogen fixation

1. Introduction

The diversity of current living forms is the result of evolutionary processes that have unfolded over billions of years but are difficult to trace in the absence of fossils. With a few exceptions, where dense sampling of natural populations can be done over short periods [1–3], only comparative genomics and phylogeny of extant species allow the prediction of ancestral gene contents and evolutionary scenarios. However, adaptive processes can be analyzed in real time via experimental evolution (EE). EE, the propagation of over tens of thousands of generations of organisms in a controlled environment [4–6], allows the investigation of the phenotypic, genomic and molecular adaptation thanks to the analysis of fossil records that are generated and stored during the experiment. When laboratory conditions mimic key conditions of natural environments, EE may allow for the understanding of natural evolutionary processes.

Nitrogen-fixing symbioses have evolved between polyphyletic soil bacteria, called rhizobia, and legumes. In the presence of compatible host plants, rhizobia elicit the formation of novel plant organs, the nodules, which they colonize massively. Inside nodule cells, bacteria reduce large amounts of nitrogen gas into ammonium, which is directly assimilated by the plant and contributes to its growth. On the plant side, this trait is predicted to result from a single gain

before the radiation of the Fabales, Fagales, Cucurbitales and Rosales clade (FaFaCuRo), which is followed by multiple independent losses in most descendant lineages [7,8], but maintained in most genera of the Fabales (legumes). On the bacterial side, the trait likely emerged in *Frankia* [9,10]. The genes that determine the synthesis of rhizobial lipo-chitooligosaccharidic Nod factors triggering nodulation are thought to have been transferred from some *Frankia* to a diazotrophic proteobacterium, generating the first rhizobium. After this, nodulation (*nod*) genes and nitrogen fixation (*nif-fix*) genes cluster together on mobile genetic elements, either plasmids or genomic islands, and spread to hundreds of species of distantly related taxa among alpha- and beta-proteobacteria via horizontal gene transfers (HGT) [11–13]. Some rhizobia, e.g., the rhizobial *Burkholderia* species, are ancient symbionts of legumes, which probably appeared shortly after legumes emerged, ca. 60 million years ago [14–16]. Others, such as *Cupriavidus taiwanensis*, the symbiont of *Mimosa pudica*, evolved more recently. *C. taiwanensis* was estimated to have emerged 12–16 million years ago following a single acquisition of a symbiotic plasmid, likely from *Burkholderia* [15,17,18]. The ancestral rhizobial *C. taiwanensis* then further diversified into five distinct genospecies with a vertical inheritance of the symbiotic plasmid in most cases (Figure 1).

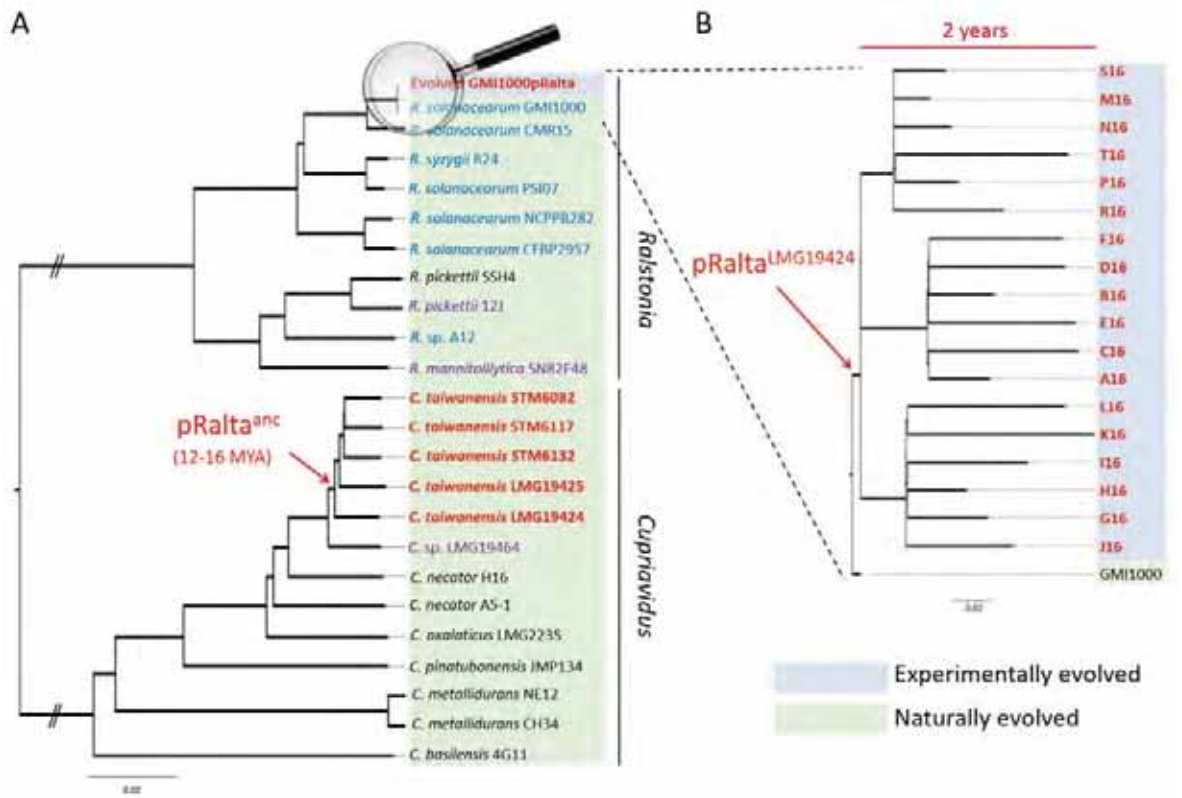


Figure 1. Phylogeny of naturally and experimentally evolved *Mimosa pudica* symbionts in the *Cupriavidus* and *Ralstonia* branches. (**A**) Neighbour-joining phylogeny of selected *Ralstonia* and *Cupriavidus* strains based on 1003 gene families defined as core genome by [17]. Evolved "GMI1000pRalta" represents a group of 18 closely related strains derived via experimental evolution from the strain *Ralstonia solanacearum* GMI1000. (**B**) Close-up of the phylogenetic relationships between the 18 resequenced clones evolved for 16 cycles (two years) compared to the ancestor (*R. solanacearum* GMI1000). This neighbour-joining phylogeny is based on single-nucleotide polymorphisms identified in the experimentally evolved strains. (A) and (B) *M. pudica* symbionts are in red and bold. Plant pathogens, opportunistic human pathogens and saprophytic strains are in blue, purple and black, respectively. Arrows indicate the acquisition of a symbiotic plasmid, either the ancestor of the *Cupriavidus taiwanensis* one (pRalta^{anc}) during natural evolution or that of *C. taiwanensis* LMG19424 (pRalta^{LMG19424}) during experimental evolution.

The *Cupriavidus* branch is a neighbour to *Ralstonia*, a genus that only contains saprophytic and pathogenic strains. The introduction of the symbiotic plasmid of *C. taiwanensis* LMG19424 into the plant pathogenic *Ralstonia solanacearum* GMI1000 strain generated a still pathogenic strain unable to nodulate *Mimosa pudica*, indicating that the acquisition of essential symbiotic genes may not be sufficient to convert a soil bacterium into a legume symbiont. To replay the emergence of a novel genus of rhizobium, here *Ralstonia*, we experimentally evolved this chimeric strain using laboratory conditions reproducing environmental conditions that may have favoured symbiosis [19,20]. Bacteria were repeatedly inoculated onto *Mimosa* plantlets in the absence of other rhizobial competitors (see Figure 2 for the experimental design). At the end of the experiment, 18 final evolved clones were isolated and their evolution analyzed (Figure 1). Despite strong differences between natural and experimental evolution of *Mimosa* symbionts in terms of genetic backgrounds, time frame and environmental conditions (Figure 1), the laboratory process exhibited striking parallels with natural evolution [17]. In particular, adaptation was characterized by a predominance of purifying selection (a purge of non-synonymous mutations) and associated with positive selection in a set of genes that led to the co-option of the same quorum-sensing system in both processes, while no adaptation was observed in the plasmid carrying the genes responsible for the ecological transition. Moreover, this experiment provided new insights into the evolutionary mechanisms leading to endosymbiosis with legumes and on the biology of rhizobia. Here we review some of the lessons learnt from our evolution experiment.

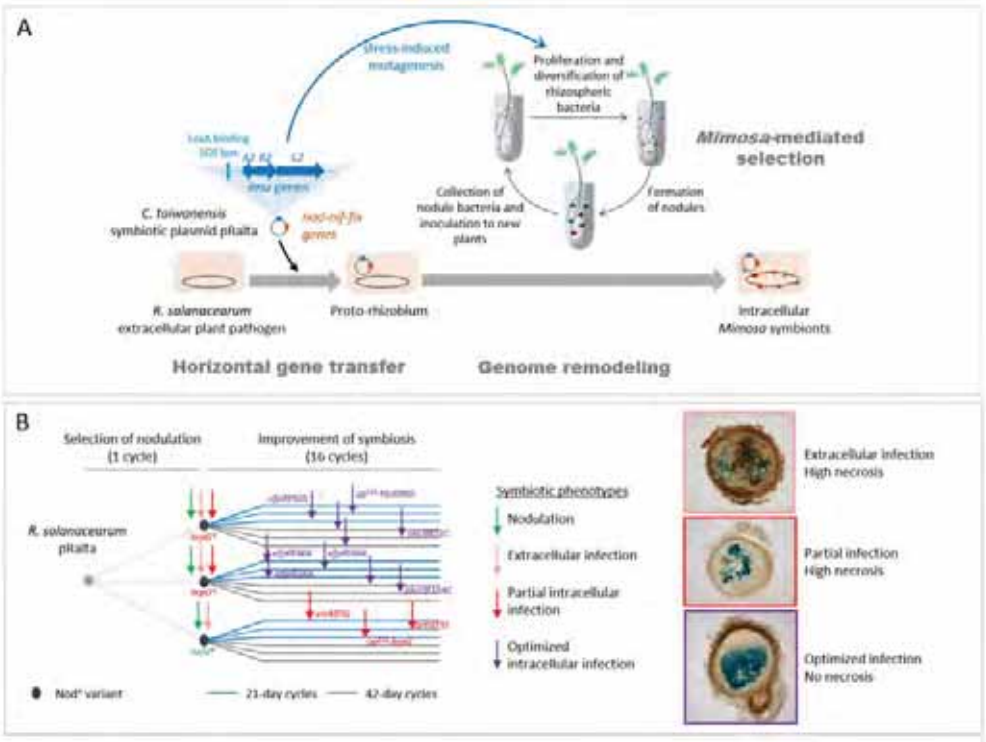


Figure 2. Experimental evolution of the plant pathogen *Ralstonia solanacearum* into legume symbionts.

(**A**) The symbiotic plasmid (pRalta) of *Cupriavidus taiwanensis*, the natural symbiont of *Mimosa pudica*, was introduced into *R. solanacearum* [21], generating a chimeric strain unable to nodulate *M. pudica*. This proto-rhizobium was then evolved through serial cycles of inoculation to *M. pudica* plantlets and re-isolation of nodule bacteria. In the first cycle, three nodulating variants were entrapped by the plant, and in the following cycles these variants were symbiotically improved (see panel B). In addition to the *nod* and *nif-fix* genes involved in Nod Factor synthesis and nitrogen fixation respectively, the pRalta plasmid carries a mutagenesis cassette (*imuA2B2C2*) that elevates mutation rate when bacteria are free-living in the plant culture medium. This transient hypermutagenesis increases the genetic diversity of the rhizospheric bacterial population among which the plant selects the most beneficial variants. (**B**) 18 parallel lineages were derived from the three nodulating (Nod^+) variants via two different selection regimes (cycles of 21 or 42 days). After acquisition of the nodulation ability, bacteria progressively improved their capacity to infect nodules. Three levels of infection were observed in the evolved lines: extracellular, partially intracellular and nicely intracellular. The main adaptive mutations responsible for the acquisition and improvement of symbiosis were identified [21–24]. *, stop mutation. FS, frameshift. up^{112}, up^{115}, intergenic mutations located 112 bp and 115 bp upstream from the gene indicated.

2. Evidence for a Two-Step Evolutionary Scenario

The pioneering works of Sullivan et al. demonstrated for the first time the conversion of soil bacteria into legume symbionts in natura through the horizontal acquisition of a symbiotic genomic island [25,26]. Since then, it has been well accepted that rhizobia diversified from independent events of horizontal transfer of essential *nod* and *nif-fix* genes. This finding was then supported by many genomic and phylogenetic analyses [11,15,27,28] and other field experiments [29,30]. However, the transferred genes are not the only ones involved in symbiosis with legumes. Many genes located outside the mobile symbiotic regions have been described as engaged in symbiosis. Among them, specific surface polysaccharide biosynthesis genes (*eps*, *kps*, *lps*, *ndv*) [31,32], metabolism genes (*glnBK*, *ilvCD*, *hemAH*, *leuCB*, *phbAC*, *dme*, *pckA*) [33–35], transporters such as *mdt*, *znu* (zinc), *dctA* (dicarboxylic acid), and *pstABC* (phosphate) genes [36–39], chaperones (*groEL*/*groES*) [40] and secretion systems [41] appear to be crucial for symbiosis in different rhizobia. Contrary to the *nod* and *nif-fix* genes, the majority of these symbiotic genes is lineage-specific [33,42]. Using in planta transposon insertion sequencing (Tn-Seq) approaches, Pobigaylo et al. and Flores-Tinoco et al. estimated that 8–15% of *Sinorhizobium meliloti* genes (500 to 900 genes) are involved in symbiosis [43,44]. These genes are scattered throughout the entire genome and mostly involved in metabolism, gene regulation and other cellular processes, confirming the diversity of genes required to achieve symbiosis. On the other hand, some bacterial functions can be deleterious for symbiosis. For example, type 3 secretion systems (T3SS) were shown to decrease the symbiotic capacity of some rhizobium strains or prevent their interaction with some host plants [41,45–47]. Furthermore, transfers of symbiotic genes performed under laboratory conditions did not always turn a non-symbiotic organism into an efficient legume symbiont. Indeed, the transfer of diverse rhizobium symbiotic plasmids into *Agrobacterium tumefaciens* or *Escherichia coli*, led to strains that were able to form either pseudo- or non-functional nodules [48–50] or nodules fixing low amounts of nitrogen [51]. In line with this, suboptimal symbionts, forming non-fixing nodules, have been observed in the wild following the horizontal acquisition of a symbiosis island [29].

We thus made the hypothesis that the evolution of a new rhizobial species/genus required a post-transfer adaptation step allowing the recruitment and/or the inactivation of specific functions of the recipient genome, especially if the transfer occurred between bacteria with different genomic backgrounds and/or lifestyles. We hypothesized that this optimization step occurred via genome remodeling under plant selection pressure. Indeed, legumes control their microsymbionts through a multistep surveillance system. All along the symbiotic process, plants regularly check the symbiotic status of bacteria and allow, or don't allow, their proliferation [52]. These checkpoints include the recognition of bacterial Nod-factors and expolysaccharides via specific receptors that convert these

signals into a plant developmental program [53–59], and the production and delivery of ammonium to plant cells, which prevent plant sanctions [60–62].

We experimentally validated this two-step evolutionary scenario (Figure 2). First, the transfer of the symbiotic plasmid of the *Mimosa* symbiont *C. taiwanensis* into the plant pathogen *Ralstonia solanacearum* led to a non-nodulating (Nod$^-$) strain, showing that, in this case, the transfer of the main symbiotic genes was not sufficient to convert *Ralstonia* into *Mimosa* symbionts. Further evolution of this Nod$^-$ *Ralstonia* chimeric strain through repeated inoculations to *Mimosa pudica* plantlets and re-isolation of bacteria from nodules allowed the parallel acquisition and progressive improvement of two symbiotic traits (nodulation and nodule cell infection) in independent lineages [19–21]. After 16 such cycles, evolved strains of most lineages induced nodules that display features of true nodules: a peripheral vascular system, the production of leghemoglobin, and cells filled with symbiosomes containing bacteria. However, mutualism was not achieved at that stage of the experiment and intracellular bacteria did not persist and prematurely degenerated. It should be noted that in nature rhizobia having suboptimal capabilities of nodulation competitiveness, infection, or nitrogen fixation, or inducing visible plant defense reactions in nodules, have been isolated [29,63,64]. As suggested by our evolution experiment, these natural isolates might correspond to intermediate evolutionary stages before the achievement of symbiosis with their host plant.

When legumes emerged, around 60–100 million years ago [14,65], nearly all extant proteobacterial genera existed and the different rhizobial lineages had already diverged [66,67]. Among α and β proteobacteria, only a few genera (15 α- and 3 β-proteobacterium genera) have been successfully colonized by symbiotic nitrogen fixation traits, suggesting that only these genera were predisposed for symbiosis. In a comparative genomics study, Garrido-Oter et al. proposed that an ancestral state of adaptation to the root environment in the Rhizobiales predisposed these bacteria to symbiosis [68]. Based on a Transposon-Sequencing (Tn-Seq) approach, Salas et al. also came to the same conclusion that rhizobia were adapted to the rhizosphere of their host plants before their specialization in nodulation [69]. The adaptation to rhizosphere colonization may include the capacity to use specific plant metabolites [70,71] and tolerance to biotic and abiotic stresses. Predisposition to legume symbiosis may require, in addition, the production of specific surface polysaccharides (EPS, KPS, LPS), and the capacity to avoid plant defenses [72,73]. Although we did not identify which genetic features predispose a bacterium to legume endosymbiosis, we provided evidence that a strain with a completely different lifestyle (pathogenic and strictly extracellular) can be converted into an intracellular legume symbiont. This revealed that indigenous functions of the recipient genome were recruited and/or suppressed for symbiosis. The capacity of *R. solanacearum* to thrive in many plant environments and infect more than 250 different plant species [74], including some legume species [75], might have contributed to its predisposition to symbiosis.

Phylogenetic studies predicted that mutualistic bacteria often derived from pathogens whilst the opposite was very rare [76]. *C. taiwanensis* might indeed have evolved from an opportunistic pathogen, since the closest identified bacterium to symbiotic *C. taiwanensis* was isolated from a cystic fibrosis patient (*C.* sp. LMG19464) [77], and these symbionts have an atypical rhizobial T3SS, which is not connected to the regulation of nodulation [45] and whose organization is similar to the T3SS of the human opportunistic bacterium *Burkholderia cenocepacia* [78]. Transitions from parasitism to mutualism were empirically demonstrated in other evolution experiments, that, for instance, converted *E. coli* parasitic phages [79], nematode pathogenic bacteria [80] or mice gut fungal pathogens [81] into beneficial symbionts.

3. Regulatory Rewiring of the Recipient Genome as a Main Driver of Symbiotic Adaptation

While the genetic mechanisms of transfer of mobile genetic elements, including symbiosis plasmids and genomic islands, have been widely studied and well described [82–85], post-transfer adaptation steps are still underexplored. Our evolution experiment allowed us to tackle this issue. To understand how *R. solanacearum* that acquired a symbiotic plasmid further evolved into *Mimosa*

intracellular symbionts, we resequenced the genome of many intermediate and final evolved clones along parallel lineages. This led us to analyze the genotype/phenotype correlations and identify the main symbiosis-adaptive mutations.

The capacity to enter roots and form nodules was unlocked by the inactivation of the T3SS (*hrcV* mutation), which is the main virulence factor of *R. solanacearum* [74]. The sole inactivation of the T3SS led to the formation of necrotic nodules that were only extracellularly invaded. a first level of intracellular infection, i.e., infection of nodule cells, was gained via the inactivation of virulence regulators, either VsrA (frameshift mutation in the *vsrA* gene) or HrpG (stop mutations in the *hrpG* gene, or mutation in the *hrpG* promoter region or in the sigma factor *prhI* acting upstream from HrpG in the regulatory cascade) [21,22]. The stop mutations in HrpG allowed the concomitant gain of the nodulation capacity since HrpG controls the T3SS and associated effectors [86], while other mutations (in *vsrA*, *hrpG* promoter and *prhI*) arose after the *hrcV* mutation having unlocked nodulation (Figure 2B). HrpG and VsrA thus likely control functions that positively or negatively interfere with intracellular infection. a small number of infected cells and the presence of necrotic zones characterized this first level of intracellular infection. a second level of infection, characterized by a high number of infected cells and an almost total absence of necrosis, was then reached in several lines. This second level of infection was obtained via mutations affecting global transcription regulators, either EfpR or PhcA. Missense mutations in *efpR* itself or an intergenic mutation upstream from a gene of unknown function triggered a constitutive repression of EfpR [23], while mutations in two different components of the Phc quorum sensing system (*phcB* and *phcQ*) led to a modulation of the quorum sensing threshold activating PhcA [24]. EfpR and PhcA were shown to act as both central players of the *R. solanacearum* virulence regulatory network and global catabolic repressors [23,87–89]. Interestingly, two genes of the *phc* regulatory pathway, *phcB* and *phcS*, display a signature of positive selection in naturally evolved *Mimosa* symbionts and we showed that PhcA was co-opted for symbiosis in *C. taiwanensis* LMG19424 [17].

Altogether, our analysis showed that the rewiring of regulatory circuits was the main driver of the symbiotic adaptation of *Ralstonia*. Strikingly, all identified adaptive mutations for infection affect global regulatory proteins (HrpG, VsrA, EfpR, PhcA) controlling, positively or negatively, the expression of several hundreds of genes (Figure 3) [23,86,89–92]. This is in accordance with other evolution experiments in which global regulators were frequently targeted [93–97]. Our results confirmed that mutations in global regulators can be highly beneficial by changing the expression of hundreds of genes in a concerted way, particularly during the first steps of adaptation to a new environment [98]. The fact that the conversion of *R. solanacearum* into intracellular legume symbionts was mainly driven by regulation rewiring means that ancillary functions required for nodule infection are there but need to be expressed at the right time and place. a future challenge will be to identify which functions controlled by the targeted regulators are beneficial or deleterious for intracellular infection in the *Ralstonia–Mimosa* interaction. We hypothesize that the same functions are affected in the different lines, but the large overlap between the EfpR and PhcA regulons [24] makes this task difficult (Figure 3).

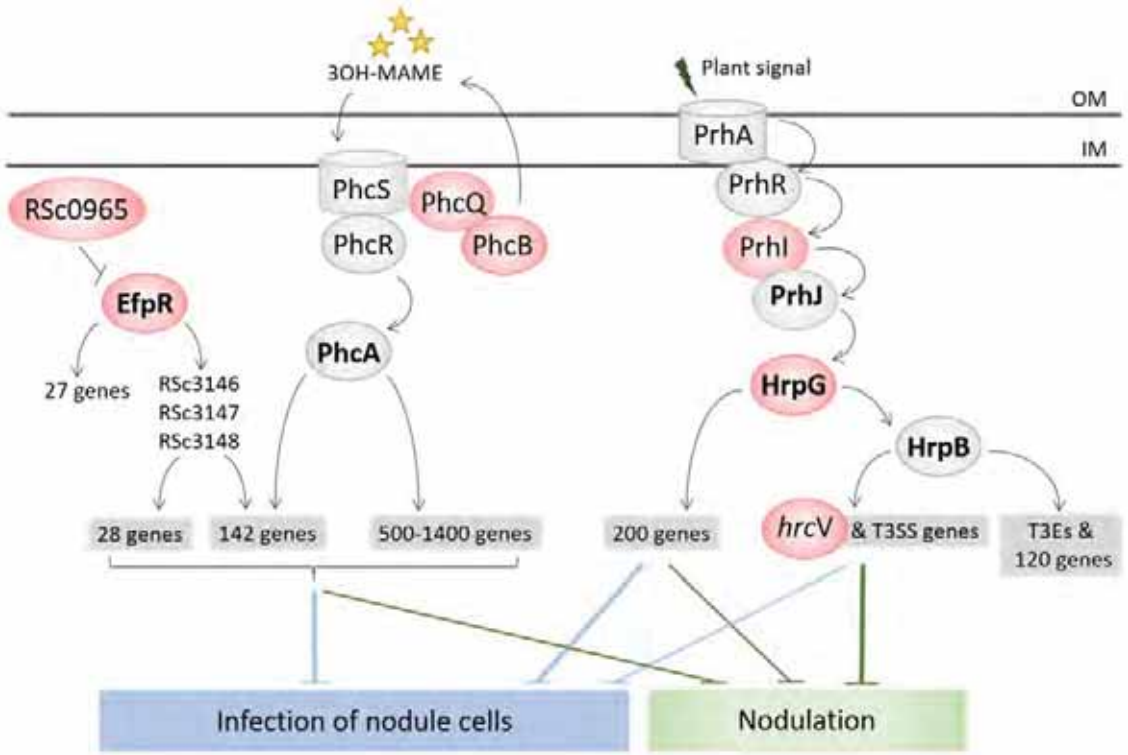


Figure 3. Adaptation to endosymbiosis mainly occurred through regulatory rewiring. Genes mutated in the experiment are indicated in red. Nodulation was gained by inactivation of the T3SS via a stop mutation in *hrcV* [21]. a first level of intracellular infection was obtained by the inactivation of the virulence regulator HrpG [21,22]. Optimized intracellular infection was reached either by inactivation of the EfpR pathway or a modulation of the quorum sensing pathway controlling PhcA [23,24]. EfpR and PhcA controls a common set of genes as well as other specific genes. It is not known which group of genes controlled by EfpR and/or PhcA interfere with intracellular infection. Transcriptional regulators are in bold. 3OH-MAME, (*R*)-3-hydroxymyristic acid methyl ester, quorum sensing molecules produced by the GMI1000 strain of *R. solanacearum*.

In natural rhizobia, a large number of regulatory systems involved in symbiosis were described. These systems vary a lot among rhizobia, even in close natural taxons. For example, the expression of nodulation genes is regulated by NodD proteins in all rhizobia. However, depending on rhizobium species, the expression of these genes is controlled by either a single NodD regulatory protein [78] or multiple, both positive and negative, regulators organized in intricate networks [99–103]. Likewise, expression of nitrogen fixation (*nif-fix*) genes is under the control of the conserved NifA protein, but upstream from NifA the regulatory cascade is highly variable among rhizobia, involving different two-component systems (FixLJ, RegSR, NtrBC, ActRS, FixL-FxkR) responding to various signals [104–107] and specific downstream transcription regulators (FixK1, FixK2, FnrN) [108–110]. Surface polysaccharide biosynthesis genes are other striking examples of essential symbiotic genes subjected to very complex regulations involving a multitude of interconnected regulators that largely differ between rhizobium species [111,112]. Integration of symbiotic functions (both endogenously recruited and horizontally acquired) in the regulatory circuits of the recipient bacterium is thus probably a key step in the evolution of new rhizobium species. In our EE, depending on the lineages, different regulatory pathways were modified by bacteria to achieve nodulation and infection, thus making a noticeable parallel between natural and experimental evolutionary processes of rhizobia.

So far, no adaptive mutations were identified on the symbiotic plasmid in the experiment, indicating that adaptation mainly occurred on the recipient genome. This is in accordance with the lack of evidence of positive selection of plasmidic genes in natural *Mimosa pudica* symbionts [17]. A possible

explanation is that in the natural evolution, like in the experimental one, the symbiotic plasmid originated from a more ancient *Mimosa* symbiont [13,16], and did not require further adaptation to the legume host.

4. The Legume Plant, a Strong Selection Pressure for Shaping Bacterial Endosymbiotic Evolution

Horizontal gene transfer (HGT) of symbiotic genes is frequent in nature [113,114], which is consistent with the observations that rhizosphere and nodule environments promote HGT [85,115]. Given that rhizobia co-exist in the soil, and even in nodules, with many other different bacteria [116,117], the transfer of symbiosis genes to new bacterial genera must be extremely common. The fact that rhizobia are restricted to a relatively small number of bacterial genera [13,118] suggests that exceptional circumstances are needed for the successful actualization of symbiotic potential in strains having received the *nod* and *nif-fix* genes. We can infer what these circumstances are by learning from field experiments with Lotus, *Biserrula pelecinus* and soybean symbionts [23,27,28,116]: (i) introduction of an exotic legume plant with a compatible rhizobium which is mal-adapted to the new environmental conditions, (ii) the presence of endogenous soil bacteria already adapted to the environment where the legume is introduced, and (iii) the absence of competing rhizobia.

Our experiment shows that, once these circumstances are reproduced in the laboratory, the plant exerts a strong selection pressure for the full expression and optimization of symbiotic traits. It was already well known that the plant selects the most nodulation-competitive strains. While large bacterial populations are present in the rhizosphere [119–121], each nodule is formed by one or a few bacteria [122]. Thus, bacterial root entry and nodulation represent a strong bottleneck and, in a mixed community, only the most competitive bacteria successfully infect the plant [123]. This property was used as a genetic screen in the past [124] and was evidenced in our EE where nodulation competitiveness strongly improved all along cycles [19]. In most cases, the genetic factors underpinning this phenotype have not been identified in our experiment (with a few exceptions, see below). Adaptive mutations improving nodulation could affect genes involved in diverse functions (growth in the rhizosphere, motility, biosynthesis of cell surface components) since it is known that multiple bacterial traits affect nodulation competitiveness [123,125].

Our evolution experiment showed for the first time that a late symbiotic stage, the infection of nodule cells, could be acquired and further improved by plant-mediated selection through two mechanisms. Firstly, bacteria multiply at much higher densities during intracellular infection (ca. 10^8–10^9 bacterial cells per nodule) than in infection threads and extracellular spaces (ca. 10^6 bacteria per nodule). Secondly, we observed a coupling between infection and nodulation since all mutations that improved intracellular infection also improved nodulation competitiveness [22] (and unpublished data). The genetic link between infection and nodulation is supported by the finding that mutations optimizing infection in our EE modify the expression of bacterial genes during the earliest stages of symbiosis (infected root hairs) [24]. This coupling is also consistent with the known effect of Nod factors on intracellular infection [126–128] and the fact that the intracellular release of rhizobia in primodium cells depends on the proper progression of infection threads [129], which otherwise conditions nodule development. From an evolutionary perspective, the nodulation/cell invasion coupling enables the selection at the root entry level of a bacterial trait that is only expressed at the later stages of the symbiotic process. These results also call for a more extensive examination of this question by systematically measuring nodulation competitiveness of non-infectious rhizobial mutants, such as exopolysaccharide deficient mutants, which has been never, or at best extremely rarely, performed.

However, the acquisition of intracellular infection was not universal in our EE as some lines (ancestor CBM356 and derived lineages with 42-day cycles) did not evolve infection (Figure 2B). Thus, depending on chance (i.e., the nature of the first nodulation-adaptive mutation) and selective regime (long (42 day) vs. short (21 day) cycles), extracellular-infecting clones may prevail over intracellular ones, possibly leading to an evolutionary dead-end.

Following their accommodation into plant cells and when oxygen concentration is low enough, rhizobia initiate nitrogen fixation, the critical process required for nutrient exchanges during this mutualistic symbiosis. a number of studies have shown that nitrogen fixation is not selected at the root entry/nodulation level per se [60,130–132], unless it is associated with another trait [133–135]. Instead, post-infection control mechanisms promote fitness of fixing bacteria over non-fixing ones [61,62]. These mechanisms were hypothesized to contribute to the stability of mutualism over long time-scales [136,137], by limiting the proliferation of non-fixing bacteria that may otherwise outcompete fixing bacteria that invest cellular resources into the energetically costly process of nitrogen fixation. In our system, we were interested in tackling another question: do post-infection control mechanisms allow the evolution of mutualism in natura and in our experimental setup? In other words, would nitrogen-fixing clones be selected in our EE, if they happen to arise? Using the reference *C. taiwanensis/M. pudica* system and isogenic Fix^+ and Fix^- bacterial mutants, we followed the dynamics of the two bacterial populations along the symbiotic process and identified the timing and extent of the post-infection control mechanisms specifically targeting Fix^- cells [60]. We show that Fix^- bacteria multiply similarly to Fix^+ bacteria in nodules, but only until 16–21 days post-infection, the time at which their fitness begins to decline compared to Fix^+ bacteria due to a premature degeneration. Interestingly, differential survival of Fix^- vs. Fix^+ bacteria occurred when single nodules were co-infected by the two bacterial genotypes. This, together with similar results obtained with *Lotus* and *Acmispon* [138], indicated the existence of cell-autonomous sanctions. Moreover, mathematical modeling predicted the impact of different environmental factors on the kinetics of invasion of rare Fix^+ clones in a Fix^- population, and these predictions were experimentally validated by witnessing the rapid increase in frequency of Fix+ bacteria along serial inoculation–nodulation cycles [60]. An important topic for future research will be to identify the molecular bases of post-infection control mechanisms. Previous work showed that soybean nodules could restrict oxygen diffusion in non-fixing nodules [61], but the generality and mechanistic details of this process remain to be investigated.

Together, our work has shown that plant-mediated selection mechanisms were efficient in promoting the successive evolution of key symbiotic properties: nodulation, intracellular infection and nitrogen fixation. While we focused here on plant-mediated selection factors observed in our EE, it is worth remembering that additional biotic and abiotic factors (including physical environment, soil bacterial communities, plant life histories) also shape bacterial adaptation in natural populations [139].

5. Discovery of a Hypermutagenesis Mechanism that Accelerates HGT-Based Evolution

The generation of genetic variation through de novo mutations fuels bacterial adaptation to new environments. Mutations in genes involved in DNA repair, such as *mutS/L/H*, can lead to a hypermutator phenotype, characterized by a constitutive elevation of mutation rate and, consequently, increased evolvability. Hypermutator genotypes are often selected in evolution experiments or during latent host colonization [140,141]. Bacteria can also increase mutation rates transiently in response to stress, a mechanism known as stress-induced mutagenesis (SIM) [142,143]. SIM is due to the expression of error-prone DNA polymerases under the control of major stress-induced regulatory pathways such as SOS or RpoS. Despite the long-term recognition of this phenomenon, the contribution of SIM to bacterial adaptive evolution has received relatively little experimental scrutiny.

Resequencing evolved clones revealed a large number of mutations in our experiment, although the strains were not constitutive hypermutators. Additional sequencing and experiments revealed that bacterial clones underwent transient hypermutagenesis (~5–10× increase in mutation rate) during each exposure to the plant culture medium, with a further increase (~4×) in the presence of *Mimosa pudica* plantlets. We identified error-prone DNA polymerases from the *imuABC* family that were located on pRalta and responsible for this transient hypermutagenesis [144]. In our system, SIM accelerates adaptation since, when re-evolved for 5 additional cycles, clones carrying *imuABC* genes (like our original ancestor) adapted faster than isogenic *imuABC* mutants evolved in the same conditions [144]. Interestingly, the interrogation of genomic databases revealed than more than half of symbiotic plasmids

carried error-prone DNA polymerases. It is tempting to speculate that the co-transfer of symbiosis genes with mutagenesis cassettes may facilitate adaptation to new symbiotic environments following plasmid acquisition by increasing the probability to generate variants with improved symbiotic capabilities that the plant can select. The fact that the natural process, as the experimental one, resulted in rapid genetic diversification dominated by purifying selection, also argues in favor of the role of *imuABC*-based transient mutagenesis in the natural evolution of rhizobia. Whether dependent on plasmidic or chromosomal genes, SIM may contribute to bursts in mutation rates occurring during host-microbe interactions [145,146]. Moreover, it is worth mentioning that other mechanisms can increase evolvability in rhizobia, such as the activation of HGT by root exudates [85], the transposition of insertion sequence elements [47,147], and the Non-Homologous End-Joining mechanism (NHEJ) [148].

Balancing the benefits during the early steps of adaptation to a new environment, hypermutagenesis can compromise fitness on the long term due to the accumulation of deleterious mutations [140]. This can lead to selection against hypermutagenesis, consistent with the modulation of mutation rates observed in several evolution experiments [149–151]. Testing whether SIM can be counter-selected would bring interesting new arguments into the debate over the adaptive nature of SIM. In that respect, it is intriguing to note that SIM does not occur in the original host of pRalta, *C. taiwanensis*, but the causes, environmental or genetic, are unknown.

6. New Insights into the Biology of Evolutionary Protagonists

EE uses extant strains with specific properties or lifestyles as protagonists. The tracking of mutations underlying phenotypic changes during laboratory evolution may reveal unidentified molecular players of functions otherwise studied by traditional molecular genetics [152,153]. In our experiment, this potential to advance the understanding of the biology of ancestral strains was highlighted by the discovery of new virulence and symbiotic components in *R. solanacearum* and *C. taiwanensis*, respectively.

The virulence mechanisms of *R. solanacearum* have been genetically dissected for many years. This allowed for the identification of a complex regulatory network involving many interconnected regulatory cascades, which controls the different functions involved in pathogenicity [154–158]. In this network, the LysR family transcriptional regulator PhcA was identified as a central regulator, whose activation is controlled by the quorum sensing (QS) PhcBSR(*Q*) system [159]. However, the role of the last gene of the operon, *phcQ,* could not be evidenced [160]. In our experiment, we focused on a *phcQ* mutation that enhanced the capacity of bacteria to infect nodule cells. This led us to show that this mutation delayed the cell density-dependent activation of PhcA and further showed that *phcQ* is involved in the production of QS molecules [24]. In addition, our experiment and another one conducted in parallel [161] revealed the occurrence of a virulence pathway previously overlooked, the *efpR* pathway. Mutations in this gene or in one upstream regulatory component allowed a better in planta colonization of *R. solanacearum* evolved on different host plants [87] and an improved intracellular infection capacity of the chimeric *Ralstonia* evolved on the *M. pudica* legume [23]. Mutation in the upstream yet unidentified regulatory component (a SNP in the intergenic region upstream from the Rsc0965 gene) led to a surprising complete loss of virulence.

The *nod* genes conveyed by symbiotic plasmids and islands determine the biosynthesis of signalling compounds, called Nod factors, which trigger the plant developmental program leading to nodulation and infection in legumes [162,163]. In 2008, Amadou et al. found that the Nod factors produced by *C. taiwanensis* were consistent with the in silico prediction based on the genome sequence and the identified *nod* genes [78]. They were found to be pentameric chito-oligomers sulphated at the reducing end, N-acylated by vaccenic acid or palmitic acid and mostly substituted by an N-methyl and a carbamoyl group at the non-reducing end. This was recently challenged from the analysis of one of our evolved lines. In the R line, the capacity to infect nodule cells was not obtained after 16 cycles of evolution despite the detection of an adaptive mutation for intracellular infection (*hrpG*A179V) in the final R16 clone [20]. Searching for the modification that countered the adaptive effect of the

*hrpG*A179V mutation led us to investigate the role of a gene deleted in the symbiotic plasmid of R16 and whose reintroduction restored nodule cell infection capacity [164]. We found that this gene, now called *noeM*, is a host-specificity nodulation gene responsible for a Nod factor modification that was previously missed, the opening and oxidation of the reducing end of the molecule [165]. Surprisingly the inactivation of *noeM* did not affect the nodule cell infection capacity but only the nodulation competitiveness of its natural host *C. taiwanensis* in interaction with *M. pudica*, indicating that its impact on symbiosis depends on the genomic background.

7. Conclusions

The evolution of nitrogen-fixing root nodule symbioses was a key innovation that gave access to new ecological niches to both plant (nitrogen-depleted soils) and bacterial (root nodules) symbionts. Comparative genomic and phylogenomic approaches have greatly contributed to understanding the evolution of these complex interactions [7,8]. By using EE on the bacterial partner, we were able to complement these approaches and provide a dynamic view of evolution, although in a somewhat artificial context, since it is impossible to know and reproduce the environmental conditions and protagonists that took part in the emergence of this symbiosis. EE did not address the question of the origin of the first rhizobia, nor that of the co-evolution between plants and rhizobia. Instead, it tackled the issue of the spreading of the N_2-fixing capacity in many proteobacterial branches. One key lesson was that a bacterium with a completely different lifestyle (pathogenic, extracellular) can be turned into an intracellular legume symbiont via the acquisition of mutualistic genes and regulatory rewiring under plant selection pressure. This result helps us to understand why so many bacterial branches have successfully been colonized by nitrogen-fixing traits. So far, our work focused on the first symbiotic stages: acquisition and improvement of nodulation and nodule invasion. By continuing the evolution experiment and its genetic analysis, future work should address the evolution of late stages of symbiosis, such as the acquisition of persistence, nitrogen fixation and mutualism and should provide a better understanding of these poorly known symbiotic stages. EE including a single selection cycle also proved to be very useful in understanding the bacterial genetic bases of incompatible symbiotic interactions [47,166,167] and to analyze plant-mediated selection acting on rhizobial communities [168,169]. Complementary to synthetic biology and classical genetic approaches, EE will undoubtedly become an indispensable tool to select efficient, host-adapted microsymbionts in the perspective of optimizing current symbiotic associations or transferring N_2-fixation in crops [170,171].

Author Contributions: G.G.D.d.M., P.R., C.M.-B., and D.C. wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the French National Research Agency (ANR-16-CE20-0011-01) and the Laboratoire d'Excellence (LABEX) TULIP (ANR-10-LABX-41). G.G.D.d.M. was supported by a fellowship from the French Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI). P.R. was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska Curie grant agreement No 845838.

Acknowledgments: We thank Finlay Warsop Thomas for careful reading of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, L.; Jelsbak, L.; Marvig, R.L.; Damkiaer, S.; Workman, C.T.; Rau, M.H.; Hansen, S.K.; Folkesson, A.; Johansen, H.K.; Ciofu, O.; et al. Evolutionary dynamics of bacteria in a human host environment. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 7481–7486. [CrossRef]
2. Zhao, S.; Lieberman, T.D.; Poyet, M.; Kauffman, K.M.; Gibbons, S.M.; Groussin, M.; Xavier, R.J.; Alm, E.J. Adaptive evolution within gut microbiomes of healthy people. *Cell Host Microbe* **2019**, *25*, 656–667. [CrossRef]
3. Garud, N.R.; Good, B.H.; Hallatschek, O.; Pollard, K.S. Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. *PLoS Biol.* **2019**, *17*, e3000102. [CrossRef] [PubMed]

4. Cooper, V.S. Experimental evolution as a high-throughput screen for genetic adaptations. *mSphere* **2018**, *3*, e00121-18. [CrossRef] [PubMed]
5. Elena, S.; Lenski, R. Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* **2003**, *4*, 457–469. [CrossRef] [PubMed]
6. Van den Bergh, B.; Swings, T.; Fauvart, M.; Michiels, J. Experimental design, population dynamics, and diversity in microbial experimental evolution. *Microbiol. Mol. Biol. Rev.* **2018**, *82*, e00008-18. [CrossRef] [PubMed]
7. Van Velzen, R.; Doyle, J.J.; Geurts, R. a resurrected scenario: Single gain and massive loss of nitrogen-fixing nodulation. *Trends Plant Sci.* **2019**, *24*, 49–57. [CrossRef] [PubMed]
8. Griesmann, M.; Chang, Y.; Liu, X.; Song, Y.; Haberer, G.; Crook, M.B.; Billault-Penneteau, B.; Lauressergues, D.; Keller, J.; Imanishi, L.; et al. Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science* **2018**, *361*, 144. [CrossRef]
9. Ktari, A.; Nouioui, I.; Furnholm, T.; Swanson, E.; Ghodhbane-Gtari, F.; Tisa, L.S.; Gtari, M. Permanent draft genome sequence of *Frankia* sp. NRRL B-16219 reveals the presence of canonical *nod* genes, which are highly homologous to those detected in *Candidatus Frankia* Dg1 genome. *Stand. Genom. Sci.* **2017**, *12*, 51. [CrossRef]
10. Persson, T.; Battenberg, K.; Demina, I.V.; Vigil-Stenman, T.; Vanden Heuvel, B.; Pujic, P.; Facciotti, M.T.; Wilbanks, E.G.; O'Brien, A.; Fournier, P.; et al. Candidatus *Frankia datiscae* Dg1, the actinobacterial microsymbiont of *Datisca glomerata*, expresses the canonical *nod* genes *nodABC* in symbiosis with its host plant. *PLoS ONE* **2015**, *10*, e0127630. [CrossRef]
11. Andrews, M.; De Meyer, S.; James, E.K.; Stępkowski, T.; Hodge, S.; Simon, M.F.; Young, J.P.W. Horizontal transfer of symbiosis genes within and between rhizobial genera: Occurrence and importance. *Genes (Basel)* **2018**, *9*, 321. [CrossRef] [PubMed]
12. MacLean, A.M.; Finan, T.M.; Sadowsky, M.J. Genomes of the symbiotic nitrogen-fixing bacteria of legumes. *Plant Physiol.* **2007**, *144*, 615–622. [CrossRef] [PubMed]
13. De Lajudie, P.; Young, J.P.W. International committee on systematics of prokaryotes subcommittee on the taxonomy of rhizobia and agrobacteria minutes of the meeting by video conference, 11 July 2018. *Int. J. Syst. Evol. Microbiol.* **2019**, *69*, 1835–1840. [CrossRef]
14. Sprent, J.I.; James, E.K. Legume evolution: Where do nodules and mycorrhizas fit in? *Plant Physiol.* **2007**, *144*, 575–581. [CrossRef] [PubMed]
15. Bontemps, C.; Elliott, G.N.; Simon, M.F.; Dos Reis Junior, F.B.; Gross, E.; Lawton, R.C.; Neto, N.E.; Loureiro, M.D.F.; De Faria, S.M.; Sprent, J.I.; et al. *Burkholderia* species are ancient symbionts of legumes. *Mol. Ecol.* **2010**, *19*, 44–52. [CrossRef] [PubMed]
16. Sprent, J.I. 60Ma of legume nodulation. What's new? What's changing? *J. Exp. Bot.* **2008**, *59*, 1081–1084. [CrossRef] [PubMed]
17. Clerissi, C.; Touchon, M.; Capela, D.; Tang, M.; Cruveiller, S.; Parker, M.A.; Moulin, L.; Masson-Boivin, C.; Rocha, E.P.C. Parallels between experimental and natural evolution of legume symbionts. *Nat. Commun.* **2018**, *9*, 2264. [CrossRef]
18. Mishra, R.P.N.; Tisseyre, P.; Melkonian, R.; Chaintreuil, C.; Miche, L.; Klonowska, A.; Gonzalez, S.; Bena, G.; Laguerre, G.; Moulin, L. Genetic diversity of *Mimosa pudica* rhizobial symbionts in soils of French Guiana: Investigating the origin and diversity of *Burkholderia phymatum* and other beta-rhizobia. *FEMS Microbiol. Ecol.* **2012**, *79*, 487–503. [CrossRef]
19. Marchetti, M.; Jauneau, A.; Capela, D.; Remigi, P.; Gris, C.; Batut, J.; Masson-Boivin, C. Shaping bacterial symbiosis with legumes by experimental evolution. *Mol. Plant-Microbe Interact.* **2014**, *27*, 956–964. [CrossRef]
20. Marchetti, M.; Clerissi, C.; Yousfi, Y.; Gris, C.; Bouchez, O.; Rocha, E.; Cruveiller, S.; Jauneau, A.; Capela, D.; Masson-Boivin, C. Experimental evolution of rhizobia may lead to either extra- or intracellular symbiotic adaptation depending on the selection regime. *Mol. Ecol.* **2017**, *26*, 1818–1831. [CrossRef]
21. Marchetti, M.; Capela, D.; Glew, M.; Cruveiller, S.; Chane-Woon-Ming, B.; Gris, C.; Timmers, T.; Poinsot, V.; Gilbert, L.B.; Heeb, P.; et al. Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biol.* **2010**, *8*, e1000280. [CrossRef] [PubMed]
22. Guan, S.H.; Gris, C.; Cruveiller, S.; Pouzet, C.; Tasse, L.; Leru, A.; Maillard, A.; Medigue, C.; Batut, J.; Masson-Boivin, C.; et al. Experimental evolution of nodule intracellular infection in legume symbionts. *ISME J.* **2013**, *7*, 1367–1377. [CrossRef] [PubMed]

23. Capela, D.; Marchetti, M.; Clérissi, C.; Perrier, A.; Guetta, D.; Gris, C.; Valls, M.; Jauneau, A.; Cruveiller, S.; Rocha, E.P.C.; et al. Recruitment of a lineage-specific virulence regulatory pathway promotes intracellular infection by a plant pathogen experimentally evolved into a legume symbiont. *Mol. Biol. Evol.* **2017**, *34*, 2503–2521. [CrossRef] [PubMed]
24. Tang, M.; Bouchez, O.; Cruveiller, S.; Masson-Boivin, C.; Capela, D. Modulation of quorum sensing as an adaptation to nodule cell infection during experimental evolution of legume symbionts. *mBio* **2020**, *11*, e03129-19. [CrossRef] [PubMed]
25. Sullivan, J.T.; Patrick, H.N.; Lowther, W.L.; Scott, D.B.; Ronson, C.W. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8985–8989. [CrossRef] [PubMed]
26. Sullivan, J.; Ronson, C. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5145–5149. [CrossRef]
27. Suominen, L.; Roos, C.; Lortet, G.; Paulin, L.; Lindström, K. Identification and structure of the *Rhizobium galegae* common nodulation genes: Evidence for horizontal gene transfer. *Mol. Biol. Evol.* **2001**, *18*, 907–916. [CrossRef]
28. Chen, W.M.; Moulin, L.; Bontemps, C.; Vandamme, P.; Bena, G.; Boivin-Masson, C. Legume symbiotic nitrogen fixation by beta-proteobacteria is widespread in nature. *J. Bacteriol.* **2003**, *185*, 7266–7272. [CrossRef]
29. Nandasena, K.G.; O'Hara, G.W.; Tiwari, R.P.; Sezmis, E.; Howieson, J.G. *In situ* lateral transfer of symbiosis islands results in rapid evolution of diverse competitive strains of mesorhizobia suboptimal in symbiotic nitrogen fixation on the pasture legume *Biserrula pelecinus* L. *Environ. Microbiol.* **2007**, *9*, 2496–2511. [CrossRef]
30. Barcellos, F.G.; Menna, P.; Batista, J.S.D.; Hungria, M. Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs *Sinorhizobium* (*Ensifer*) *fredii* and *Bradyrhizobium elkanii* in a Brazilian Savannah soil. *Appl. Environ. Microbiol.* **2007**, *73*, 2635–2643. [CrossRef]
31. Downie, J.A. The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. *FEMS Microbiol. Rev.* **2010**, *34*, 150–170. [CrossRef] [PubMed]
32. Fraysse, N.; Couderc, F.; Poinsot, V. Surface polysaccharide involvement in establishing the rhizobium-legume symbiosis. *Eur. J. Biochem.* **2003**, *270*, 1365–1380. [CrossRef]
33. Tian, C.F.; Zhou, Y.J.; Zhang, Y.M.; Li, Q.Q.; Zhang, Y.Z.; Li, D.F.; Wang, S.; Wang, J.; Gilbert, L.B.; Li, Y.R.; et al. Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 8629–8634. [CrossRef] [PubMed]
34. Zhang, Y.; Aono, T.; Poole, P.; Finan, T.M. NAD(P)+-malic enzyme mutants of *Sinorhizobium* sp. strain NGR234, but not *Azorhizobium caulinodans* ORS571, maintain symbiotic N_2 fixation capabilities. *Appl. Environ. Microbiol.* **2012**, *78*, 2803–2812. [CrossRef] [PubMed]
35. Yurgel, S.N.; Rice, J.; Kahn, M.L. Transcriptome analysis of the role of GlnD/GlnBK in nitrogen stress adaptation by *Sinorhizobium meliloti* Rm1021. *PLoS ONE* **2013**, *8*, e58028. [CrossRef] [PubMed]
36. Batista, S.; Catalán, A.I.; Hernández-Lucas, I.; Martínez-Romero, E.; Aguilar, O.M.; Martínez-Drets, G. Identification of a system that allows a *Rhizobium tropici dctA* mutant to grow on succinate, but not on other C4-dicarboxylates. *Can. J. Microbiol.* **2001**, *47*, 509–518. [CrossRef]
37. Jiao, J.; Wu, L.J.; Zhang, B.; Hu, Y.; Li, Y.; Zhang, X.X.; Guo, H.J.; Liu, L.X.; Chen, W.X.; Zhang, Z.; et al. MucR is required for transcriptional activation of conserved ion transporters to support nitrogen fixation of *Sinorhizobium fredii* in soybean nodules. *Mol. Plant Microbe Interact.* **2016**, *29*, 352–361. [CrossRef]
38. Hu, Y.; Jiao, J.; Liu, L.X.; Sun, Y.W.; Chen, W.F.; Sui, X.H.; Chen, W.X.; Tian, C.F. Evidence for phosphate starvation of rhizobia without terminal differentiation in legume nodules. *Mol. Plant Microbe Interact.* **2018**, *31*, 1060–1068. [CrossRef]
39. Jiao, J.; Ni, M.; Zhang, B.; Zhang, Z.; Young, J.P.W.; Chan, T.F.; Chen, W.X.; Lam, H.M.; Tian, C.F. Coordinated regulation of core and accessory genes in the multipartite genome of *Sinorhizobium fredii*. *PLoS Genet.* **2018**, *14*, e1007428. [CrossRef]
40. Bittner, A.N.; Foltz, A.; Oke, V. Only one of five *groEL* genes is required for viability and successful symbiosis in *Sinorhizobium meliloti*. *J. Bacteriol.* **2007**, *189*, 1884–1889. [CrossRef]
41. Nelson, M.S.; Sadowsky, M.J. Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. *Front. Plant Sci.* **2015**, *6*, 491. [CrossRef] [PubMed]
42. Black, M.; Moolhuijzen, P.; Chapman, B.; Barrero, R.; Howieson, J.; Hungria, M.; Bellgard, M. The genetics of symbiotic nitrogen fixation: Comparative genomics of 14 rhizobia strains by resolution of protein clusters. *Genes (Basel)* **2012**, *3*, 138–166. [CrossRef] [PubMed]

43. Pobigaylo, N.; Szymczak, S.; Nattkemper, T.W.; Becker, A. Identification of genes relevant to symbiosis and competitiveness in *Sinorhizobium meliloti* using signature-tagged mutants. *Mol. Plant Microbe Interact.* **2008**, *21*, 219–231. [CrossRef] [PubMed]
44. Flores-Tinoco, C.E.; Christen, M.; Christen, B. Co-catabolism of arginine and succinate drives symbiotic nitrogen fixation. *BioRxiv* **2019**. [CrossRef]
45. Saad, M.M.; Crèvecoeur, M.; Masson-Boivin, C.; Perret, X. The type 3 protein secretion system of *Cupriavidus taiwanensis* strain LMG19424 compromises symbiosis with *Leucaena leucocephala*. *Appl. Environ. Microbiol.* **2012**, *78*, 7476–7479. [CrossRef]
46. Viprey, V.; Del Greco, A.; Golinowski, W.; Broughton, W.J.; Perret, X. Symbiotic implications of type III protein secretion machinery in Rhizobium. *Mol. Microbiol.* **1998**, *28*, 1381–1389. [CrossRef]
47. Zhao, R.; Liu, L.X.; Zhang, Y.Z.; Jiao, J.; Cui, W.J.; Zhang, B.; Wang, X.L.; Li, M.L.; Chen, Y.; Xiong, Z.Q.; et al. Adaptive evolution of rhizobial symbiotic compatibility mediated by co-evolved insertion sequences. *ISME J.* **2018**, *12*, 101–111. [CrossRef]
48. Hirsch, A.M.; Wilson, K.J.; Jones, J.D.; Bang, M.; Walker, V.V.; Ausubel, F.M. *Rhizobium meliloti* nodulation genes allow *Agrobacterium tumefaciens* and *Escherichia coli* to form pseudonodules on alfalfa. *J. Bacteriol.* **1984**, *158*, 1133–1143. [CrossRef]
49. Abe, M.; Kawamura, R.; Higashi, S.; Mori, S.; Shibata, M.; Uchiumi, T. Transfer of the symbiotic plasmid from *Rhizobium leguminosarum* biovar *trifolii* to *Agrobacterium tumefaciens*. *J. Gen. Appl. Microbiol.* **1998**, *44*, 65–74. [CrossRef]
50. Nakatsukasa, H.; Uchiumi, T.; Kucho, K.; Suzuki, A.; Higashi, S.; Abe, M. Transposon mediation allows a symbiotic plasmid of *Rhizobium leguminosarum* bv. *trifolii* to become a symbiosis island in *Agrobacterium* and *Rhizobium*. *J. Gen. Appl. Microbiol.* **2008**, *54*, 107–118. [CrossRef]
51. Martinez, E.; Palacios, R.; Sanchez, F. Nitrogen fixing nodules induced by *Agrobacterium tumefaciens* harboring *Rhizobium phaseoli* plasmids. *J. Bacteriol.* **1987**, *169*, 2828–2834. [CrossRef] [PubMed]
52. Benezech, C.; Doudement, M.; Gourion, B. Legumes tolerance to rhizobia is not always observed and not always deserved. *Cell. Microbiol.* **2020**, *22*, e13124. [CrossRef] [PubMed]
53. Limpens, E.; Franken, C.; Smit, P.; Willemse, J.; Bisseling, T.; Geurts, R. LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* **2003**, *302*, 630–633. [CrossRef] [PubMed]
54. Madsen, E.B.; Madsen, L.H.; Radutoiu, S.; Olbryt, M.; Rakwalska, M.; Szczyglowski, K.; Sato, S.; Kaneko, T.; Tabata, S.; Sandal, N.; et al. a receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* **2003**, *425*, 637–640. [CrossRef]
55. Radutoiu, S.; Madsen, L.H.; Madsen, E.B.; Felle, H.H.; Umehara, Y.; Grønlund, M.; Sato, S.; Nakamura, Y.; Tabata, S.; Sandal, N.; et al. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **2003**, *425*, 585–592. [CrossRef]
56. 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.; et al. The *Medicago truncatula* lysin motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant Physiol.* **2006**, *142*, 265–279. [CrossRef]
57. Mulder, L.; Lefebvre, B.; Cullimore, J.; Imberty, A. LysM domains of *Medicago truncatula* NFP protein involved in Nod factor perception. Glycosylation state, molecular modeling and docking of chitooligosaccharides and Nod factors. *Glycobiology* **2006**, *16*, 801–809. [CrossRef]
58. Smit, P.; Limpens, E.; Geurts, R.; Fedorova, E.; Dolgikh, E.; Gough, C.; Bisseling, T. *Medicago* LYK3, an entry receptor in rhizobial nodulation factor signaling. *Plant Physiol.* **2007**, *145*, 183–191. [CrossRef]
59. 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.; et al. Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature* **2015**, *523*, 308–312. [CrossRef]
60. Daubech, B.; Remigi, P.; Doin de Moura, G.; Marchetti, M.; Pouzet, C.; Auriac, M.C.; Gokhale, C.S.; Masson-Boivin, C.; Capela, D. Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation. *Elife* **2017**, *6*. [CrossRef]
61. Kiers, E.; Rousseau, R.; West, S.; Denison, R. Host sanctions and the legume-rhizobium mutualism. *Nature* **2003**, *425*, 78–81. [CrossRef] [PubMed]
62. Oono, R.; Anderson, C.G.; Denison, R.F. Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. *Proc. Biol. Sci.* **2011**, *278*, 2698–2703. [CrossRef]

63. Gehlot, H.S.; Tak, N.; Kaushik, M.; Mitra, S.; Chen, W.-M.; Poweleit, N.; Panwar, D.; Poonar, N.; Parihar, R.; Tak, A.; et al. An invasive *Mimosa* in India does not adopt the symbiont of its native relatives. *Ann. Bot.* **2013**, *112*, 179–196. [CrossRef] [PubMed]
64. Gossmann, J.A.; Markmann, K.; Brachmann, A.; Rose, L.E.; Parniske, M. Polymorphic infection and organogenesis patterns induced by a *Rhizobium leguminosarum* isolate from *Lotus* root nodules are determined by the host genotype. *New Phytol.* **2012**, *196*, 561–573. [CrossRef] [PubMed]
65. Werner, G.D.A.; Cornwell, W.K.; Sprent, J.I.; Kattge, J.; Kiers, E.T. a single evolutionary innovation drives the deep evolution of symbiotic N_2-fixation in angiosperms. *Nat. Commun.* **2014**, *5*, 4087. [CrossRef]
66. Turner, S.L.; Young, J.P.W. The glutamine synthetases of rhizobia: Phylogenetics and evolutionary implications. *Mol. Biol. Evol.* **2000**, *17*, 309–319. [CrossRef]
67. Marin, J.; Battistuzzi, F.U.; Brown, A.C.; Hedges, S.B. The timetree of prokaryotes: New insights into their evolution and speciation. *Mol. Biol. Evol.* **2017**, *34*, 437–446. [CrossRef]
68. Garrido-Oter, R.; Nakano, R.T.; Dombrowski, N.; Ma, K.W.; McHardy, A.C.; Schulze-Lefert, P.; Team, A. Modular traits of the Rhizobiales root microbiota and their evolutionary relationship with symbiotic rhizobia. *Cell Host Microbe* **2018**, *24*, 155–167. [CrossRef]
69. Salas, M.E.; Lozano, M.J.; López, J.L.; Draghi, W.O.; Serrania, J.; Torres Tejerizo, G.A.; Albicoro, F.J.; Nilsson, J.F.; Pistorio, M.; Del Papa, M.F.; et al. Specificity traits consistent with legume-rhizobia coevolution displayed by *Ensifer meliloti* rhizosphere colonization. *Environ. Microbiol.* **2017**, *19*, 3423–3438. [CrossRef]
70. Ding, H.; Yip, C.B.; Geddes, B.A.; Oresnik, I.J.; Hynes, M.F. Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. *Microbiology* **2012**, *158*, 1369–1378. [CrossRef]
71. Garcia-Fraile, P.; Seaman, J.C.; Karunakaran, R.; Edwards, A.; Poole, P.S.; Downie, J.A. Arabinose and protocatechuate catabolism genes are important for growth of *Rhizobium leguminosarum* biovar *viciae* in the pea rhizosphere. *Plant Soil* **2015**, *390*, 251–264. [CrossRef] [PubMed]
72. Gourion, B.; Berrabah, F.; Ratet, P.; Stacey, G. Rhizobium-legume symbioses: The crucial role of plant immunity. *Trends Plant Sci.* **2015**, *20*, 186–194. [CrossRef] [PubMed]
73. Berrabah, F.; Bourcy, M.; Eschstruth, A.; Cayrel, A.; Guefrachi, I.; Mergaert, P.; Wen, J.; Jean, V.; Mysore, K.S.; Gourion, B.; et al. a nonRD receptor-like kinase prevents nodule early senescence and defense-like reactions during symbiosis. *New Phytol.* **2014**, *203*, 1305–1314. [CrossRef] [PubMed]
74. Genin, S.; Denny, T.P. Pathogenomics of the *Ralstonia solanacearum* species complex. *Ann. Rev. Phytopathol.* **2012**, *50*, 67–89. [CrossRef] [PubMed]
75. Vailleau, F.; Sartorel, E.; Jardinaud, M.F.; Chardon, F.; Genin, S.; Huguet, T.; Gentzbittel, L.; Petitprez, M. Characterization of the interaction between the bacterial wilt pathogen *Ralstonia solanacearum* and the model legume plant *Medicago truncatula*. *Mol. Plant Microbe Interact.* **2007**, *20*, 159–167. [CrossRef] [PubMed]
76. Sachs, J.L.; Skophammer, R.G.; Bansal, N.; Stajich, J.E. Evolutionary origins and diversification of proteobacterial mutualists. *Proc. Biol. Sci.* **2014**, *281*, 20132146. [CrossRef]
77. Chen, W.M.; Laevens, S.; Lee, T.M.; Coenye, T.; De Vos, P.; Mergeay, M.; Vandamme, P. *Ralstonia taiwanensis* sp nov., isolated from root nodules of *Mimosa* species and sputum of a cystic fibrosis patient. *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 1729–1735. [CrossRef]
78. Amadou, C.; Pascal, G.; Mangenot, S.; Glew, M.; Bontemps, C.; Capela, D.; Carrere, S.; Cruveiller, S.; Dossat, C.; Lajus, A.; et al. Genome sequence of the beta-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Res.* **2008**, *18*, 1472–1483. [CrossRef]
79. Shapiro, J.W.; Turner, P.E. Evolution of mutualism from parasitism in experimental virus populations. *Evolution* **2018**, *72*, 707–712. [CrossRef]
80. King, K.C.; Brockhurst, M.A.; Vasieva, O.; Paterson, S.; Betts, A.; Ford, S.A.; Frost, C.L.; Horsburgh, M.J.; Haldenby, S.; Hurst, G.D. Rapid evolution of microbe-mediated protection against pathogens in a worm host. *ISME J.* **2016**, *10*, 1915–1924. [CrossRef]
81. Tso, G.H.W.; Reales-Calderon, J.A.; Tan, A.S.M.; Sem, X.; Le, G.T.T.; Tan, T.G.; Lai, G.C.; Srinivasan, K.G.; Yurieva, M.; Liao, W.; et al. Experimental evolution of a fungal pathogen into a gut symbiont. *Science* **2018**, *362*, 589–595. [CrossRef] [PubMed]
82. Bañuelos-Vazquez, L.A.; Torres Tejerizo, G.; Brom, S. Regulation of conjugative transfer of plasmids and integrative conjugative elements. *Plasmid* **2017**, *91*, 82–89. [CrossRef] [PubMed]

83. Ding, H.; Hynes, M.F. Plasmid transfer systems in the rhizobia. *Can. J. Microbiol.* **2009**, *55*, 917–927. [CrossRef] [PubMed]
84. Ramsay, J.P.; Tester, L.G.L.; Major, A.S.; Sullivan, J.T.; Edgar, C.D.; Kleffmann, T.; Patterson-House, J.R.; Hall, D.A.; Tate, W.P.; Hynes, M.F.; et al. Ribosomal frameshifting and dual-target antiactivation restrict quorum-sensing-activated transfer of a mobile genetic element. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 4104–4109. [CrossRef] [PubMed]
85. Ling, J.; Wang, H.; Wu, P.; Li, T.; Tang, Y.; Naseer, N.; Zheng, H.; Masson-Boivin, C.; Zhong, Z.; Zhu, J. Plant nodulation inducers enhance horizontal gene transfer of *Azorhizobium caulinodans* symbiosis island. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13875–13880. [CrossRef] [PubMed]
86. Valls, M.; Genin, S.; Boucher, C. Integrated regulation of the type III secretion system and other virulence determinants in *Ralstonia solanacearum*. *PLoS Pathog.* **2006**, *2*, 798–807. [CrossRef]
87. Perrier, A.; Peyraud, R.; Rengel, D.; Barlet, X.; Lucasson, E.; Gouzy, J.; Peeters, N.; Genin, S.; Guidot, A. Enhanced in planta fitness through adaptive mutations in *efpR*, a dual regulator of virulence and metabolic functions in the plant pathogen *Ralstonia solanacearum*. *PLoS Pathog.* **2016**, *12*, e1006044. [CrossRef]
88. Peyraud, R.; Cottret, L.; Marmiesse, L.; Gouzy, J.; Genin, S. a Resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. *PLoS Pathog.* **2016**, *12*, e1005939. [CrossRef]
89. Perrier, A.; Barlet, X.; Peyraud, R.; Rengel, D.; Guidot, A.; Genin, S. Comparative transcriptomic studies identify specific expression patterns of virulence factors under the control of the master regulator PhcA in the *Ralstonia solanacearum* species complex. *Microb. Pathog.* **2018**, *116*, 273–278. [CrossRef]
90. Mori, Y.; Ishikawa, S.; Ohnishi, H.; Shimatani, M.; Morikawa, Y.; Hayashi, K.; Ohnishi, K.; Kiba, A.; Kai, K.; Hikichi, Y. Involvement of ralfuranones in the quorum sensing signalling pathway and virulence of *Ralstonia solanacearum* strain OE1-1. *Mol. Plant Pathol.* **2018**, *19*, 454–463. [CrossRef]
91. Perrier, A.; Barlet, X.; Rengel, D.; Prior, P.; Poussier, S.; Genin, S.; Guidot, A. Spontaneous mutations in a regulatory gene induce phenotypic heterogeneity and adaptation of *Ralstonia solanacearum* to changing environments. *Environ. Microbiol.* **2019**, *21*, 3140–3152. [CrossRef] [PubMed]
92. Khokhani, D.; Lowe-Power, T.M.; Tran, T.M.; Allen, C. a Single regulator mediates strategic switching between attachment/spread and growth/virulence in the plant pathogen. *mBio* **2017**, *8*, e00895-17. [CrossRef] [PubMed]
93. Philippe, N.; Crozat, E.; Lenski, R.E.; Schneider, D. Evolution of global regulatory networks during a long-term experiment with *Escherichia coli*. *Bioessays* **2007**, *29*, 846–860. [CrossRef] [PubMed]
94. Hindré, T.; Knibbe, C.; Beslon, G.; Schneider, D. New insights into bacterial adaptation through *in vivo* and *in silico* experimental evolution. *Nat. Rev. Microbiol.* **2012**, *10*, 352–365. [CrossRef]
95. Pankey, S.M.; Foxall, R.L.; Ster, I.M.; Perry, L.A.; Schuster, B.M.; Donner, R.A.; Coyle, M.; Cooper, V.S.; Whistler, C.A. Host-selected mutations converging on a global regulator drive an adaptive leap towards symbiosis in bacteria. *Elife* **2017**, *6*, e24414. [CrossRef]
96. Carroll, S.M.; Chubiz, L.M.; Agashe, D.; Marx, C.J. Parallel and divergent evolutionary solutions for the optimization of an engineered central metabolism in *Methylobacterium extorquens* AM1. *Microorganisms* **2015**, *3*, 152–174. [CrossRef]
97. Damkiær, S.; Yang, L.; Molin, S.; Jelsbak, L. Evolutionary remodeling of global regulatory networks during long-term bacterial adaptation to human hosts. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7766–7771. [CrossRef]
98. Rodriguez-Verdugo, A.; Tenaillon, O.; Gaut, B.S. First-Step Mutations during adaptation restore the expression of hundreds of genes. *Mol. Biol. Evol.* **2016**, *33*, 25–39. [CrossRef]
99. 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]
100. Del Cerro, P.; Rolla-Santos, A.A.; Gomes, D.F.; Marks, B.B.; del Rosario Espuny, M.; Rodríguez-Carvajal, M.; Soria-Díaz, M.E.; Nakatani, A.S.; Hungria, M.; Ollero, F.J.; et al. Opening the black box of *nodD3*, *nodD4* and *nodD5* genes of *Rhizobium tropici* strain CIAT 899. *BMC Genom.* **2015**, *16*, 864. [CrossRef]
101. Kobayashi, H.; Graven, Y.N.; Broughton, W.J.; Perret, X. Flavonoids induce temporal shifts in gene-expression of *nod*-box controlled loci in *Rhizobium* sp NGR234. *Mol. Microbiol.* **2004**, *51*, 335–347. [CrossRef] [PubMed]
102. Chen, H.; Gao, K.; Kondorosi, E.; Kondorosi, A.; Rolfe, B.G. Functional genomic analysis of global regulator NolR in *Sinorhizobium meliloti*. *Mol. Plant Microbe Interact.* **2005**, *18*, 1340–1352. [CrossRef] [PubMed]

103. Del Cerro, P.; Rolla-Santos, A.A.; Valderrama-Fernández, R.; Gil-Serrano, A.; Bellogín, R.A.; Gomes, D.F.; Pérez-Montaño, F.; Megías, M.; Hungría, M.; Ollero, F.J. NrcR, a New transcriptional regulator of *Rhizobium tropici* CIAT 899 involved in the legume root-nodule symbiosis. *PLoS ONE* **2016**, *11*, e0154029. [CrossRef] [PubMed]
104. Bauer, E.; Kaspar, T.; Fischer, H.M.; Hennecke, H. Expression of the *fixR-nifA* operon in *Bradyrhizobium japonicum* depends on a new response regulator, RegR. *J. Bacteriol.* **1998**, *180*, 3853–3863. [CrossRef] [PubMed]
105. Kaminski, P.A.; Elmerich, C. The control of *Azorhizobium caulinodans nifA* expression by oxygen, ammonia and by the HF-I-like protein, NrfA. *Mol. Microbiol.* **1998**, *28*, 603–613. [CrossRef]
106. Cosseau, C.; Batut, J. Genomics of the *ccoNOQP*-encoded *cbb*(3) oxidase complex in bacteria. *Arch. Microbiol.* **2004**, *181*, 89–96. [CrossRef]
107. Fenner, B.J.; Tiwari, R.P.; Reeve, W.G.; Dilworth, M.J.; Glenn, A.R. *Sinorhizobium medicae* genes whose regulation involves the ActS and/or ActR signal transduction proteins. *FEMS Microbiol. Lett.* **2004**, *236*, 21–31. [CrossRef]
108. Moris, M.; Dombrecht, B.; Xi, C.; Vanderleyden, J.; Michiels, J. Regulatory role of *Rhizobium etli* CNPAF512 *fnrN* during symbiosis. *Appl. Environ. Microbiol.* **2004**, *70*, 1287–1296. [CrossRef]
109. Reyes-González, A.; Talbi, C.; Rodríguez, S.; Rivera, P.; Zamorano-Sánchez, D.; Girard, L. Expanding the regulatory network that controls nitrogen fixation in *Sinorhizobium meliloti*: Elucidating the role of the two-component system hFixL-FxkR. *Microbiology* **2016**, *162*, 979–988. [CrossRef]
110. Mesa, S.; Hauser, F.; Friberg, M.; Malaguti, E.; Fischer, H.M.; Hennecke, H. Comprehensive assessment of the regulons controlled by the FixLJ-FixK2-FixK1 cascade in *Bradyrhizobium japonicum*. *J. Bacteriol.* **2008**, *190*, 6568–6579. [CrossRef]
111. Janczarek, M. Environmental signals and regulatory pathways that influence exopolysaccharide production in rhizobia. *Int. J. Mol. Sci.* **2011**, *12*, 7898–7933. [CrossRef] [PubMed]
112. Bonomi, H.R.; Posadas, D.M.; Paris, G.; Carrica, M.e.C.; Frederickson, M.; Pietrasanta, L.I.; Bogomolni, R.A.; Zorreguieta, A.; Goldbaum, F.A. Light regulates attachment, exopolysaccharide production, and nodulation in *Rhizobium leguminosarum* through a LOV-histidine kinase photoreceptor. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 12135–12140. [CrossRef]
113. Lemaire, B.; Van Cauwenberghe, J.; Chimphango, S.; Stirton, C.; Honnay, O.; Smets, E.; Muasya, A.M. Recombination and horizontal transfer of nodulation and ACC deaminase (*acdS*) genes within Alpha- and Betaproteobacteria nodulating legumes of the Cape Fynbos biome. *FEMS Microbiol. Ecol.* **2015**, *91*, fiv118. [CrossRef]
114. Pérez Carrascal, O.M.; VanInsberghe, D.; Juárez, S.; Polz, M.F.; Vinuesa, P.; González, V. Population genomics of the symbiotic plasmids of sympatric nitrogen-fixing *Rhizobium* species associated with *Phaseolus vulgaris*. *Environ. Microbiol.* **2016**, *18*, 2660–2676. [CrossRef] [PubMed]
115. Bañuelos-Vazquez, L.A.; Torres Tejerizo, G.; Cervantes-De La Luz, L.; Girard, L.; Romero, D.; Brom, S. Conjugative transfer between *Rhizobium etli* endosymbionts inside the root nodule. *Environ. Microbiol.* **2019**, *21*, 3430–3441. [CrossRef] [PubMed]
116. Zgadzaj, R.; James, E.K.; Kelly, S.; Kawaharada, Y.; de Jonge, N.; Jensen, D.B.; Madsen, L.H.; Radutoiu, S. a Legume genetic framework controls infection of nodules by symbiotic and endophytic bacteria. *PLoS Genet.* **2015**, *11*, e1005280. [CrossRef] [PubMed]
117. Martínez-Hidalgo, P.; Hirsch, A. The nodule microbiome: N_2-fixing rhizobia do not live alone. In *Phytobiomes Journal*; Carolyn, A., Young, C.A., Eds.; The American Phytopathological Society: St. Paul, MN, USA, 2017; Volume 1, pp. 70–82. [CrossRef]
118. Tang, M.; Capela, D. Rhizobium diversity in the light of evolution. In *Advances in Botanic Research: Regulations of Nitrogen-Fixing Symbioses in Legumes*; Frendo, P., Frugier, F., Masson-Boivin, C., Eds.; Elsevier: Amsterdam, The Netherlands, 2020; Volume 94, pp. 251–288.
119. Zgadzaj, R.; Garrido-Oter, R.; Jensen, D.B.; Koprivova, A.; Schulze-Lefert, P.; Radutoiu, S. Root nodule symbiosis in *Lotus japonicus* drives the establishment of distinctive rhizosphere, root, and nodule bacterial communities. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E7996–E8005. [CrossRef]
120. Mendes, R.; Garbeva, P.; Raaijmakers, J.M. The rhizosphere microbiome: Significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol. Rev.* **2013**, *37*, 634–663. [CrossRef]

121. Xiao, X.; Chen, W.; Zong, L.; Yang, J.; Jiao, S.; Lin, Y.; Wang, E.; Wei, G. Two cultivated legume plants reveal the enrichment process of the microbiome in the rhizocompartments. *Mol. Ecol.* **2017**, *26*, 1641–1651. [CrossRef]
122. Gage, D.J. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 280–300. [CrossRef]
123. Triplett, E.W.; Sadowsky, M.J. Genetics of competition for nodulation of Legumes. *Ann. Rev. Microbiol.* **1992**, *46*, 399–428. [CrossRef] [PubMed]
124. Long, S.R.; Buikema, W.J.; Ausubel, F.M. Cloning of *Rhizobium meliloti* nodulation genes by direct complementation of Nod$^-$ mutants. *Nature* **1982**, *298*, 485–488. [CrossRef]
125. Clúa, J.; Roda, C.; Zanetti, M.E.; Blanco, F.A. Compatibility between legumes and rhizobia for the establishment of a successful nitrogen-fixing symbiosis. *Genes (Basel)* **2018**, *9*, 125. [CrossRef] [PubMed]
126. Capoen, W.; Goormachtig, S.; De Rycke, R.; Schroeyers, K.; Holsters, M. SrSymRK, a plant receptor essential for symbiosome formation. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10369–10374. [CrossRef]
127. Ovchinnikova, E.; Journet, E.-P.; Chabaud, M.; Cosson, V.; Ratet, P.; Duc, G.; Fedorova, E.; Liu, W.; den Camp, R.O.; Zhukov, V.; et al. IPD3 controls the formation of nitrogen-fixing symbiosomes in Pea and *Medicago* Spp. *Mol. Plant-Microbe Interact.* **2011**, *24*, 1333–1344. [CrossRef]
128. Moling, S.; Pietraszewska-Bogiel, A.; Postma, M.; Fedorova, E.; Hink, M.A.; Limpens, E.; Gadella, T.W.; Bisseling, T. Nod factor receptors form heteromeric complexes and are essential for intracellular infection in *Medicago* nodules. *Plant Cell* **2014**, *26*, 4188–4199. [CrossRef]
129. Xiao, T.T.; Schilderink, S.; Moling, S.; Deinum, E.E.; Kondorosi, E.; Franssen, H.; Kulikova, O.; Niebel, A.; Bisseling, T. Fate map of *Medicago truncatula* root nodules. *Development* **2014**, *141*, 3517–3528. [CrossRef]
130. Westhoek, A.; Field, E.; Rehling, F.; Mulley, G.; Webb, I.; Poole, P.S.; Turnbull, L.A. Policing the legume-*Rhizobium* symbiosis: a critical test of partner choice. *Sci. Rep.* **2017**, *7*, 1419. [CrossRef]
131. Hahn, M.; Studer, D. Competitiveness of a *nif Bradyrhizobium japonicum* mutant against the wild-type strain. *FEMS Microbiol. Lett.* **1986**, *33*, 143–148. [CrossRef]
132. Amarger, N. Competition for nodule formation between effective and ineffective strains of *Rhizobium meliloti*. *Soil Biol. Biochem.* **1981**, *13*, 475–480. [CrossRef]
133. Gubry-Rangin, C.; Garcia, M.; Bena, G. Partner choice in *Medicago Truncatula-Sinorhizobium* symbiosis. *Proc. Biol. Sci.* **2010**, *277*, 1947–1951. [CrossRef] [PubMed]
134. Heath, K.D.; Tiffin, P. Stabilizing mechanisms in a legume-rhizobium mutualism. *Evolution* **2009**, *63*, 652–662. [CrossRef] [PubMed]
135. Younginger, B.S.; Friesen, M.L. Connecting signals and benefits through partner choice in plant-microbe interactions. *FEMS Microbiol. Lett.* **2019**, *366*, fnz217. [CrossRef] [PubMed]
136. Kiers, E.T.; Denison, R.F. Sanctions, cooperation, and the stability of plant-rhizosphere mutualisms. *Ann. Rev. Ecol. Evol. Syst.* **2008**, *39*, 215–236. [CrossRef]
137. Friesen, M.L. Widespread fitness alignment in the legume-rhizobium symbiosis. *New Phytol.* **2012**, *194*, 1096–1111. [CrossRef] [PubMed]
138. Regus, J.U.; Quides, K.W.; O'Neill, M.R.; Suzuki, R.; Savory, E.A.; Chang, J.H.; Sachs, J.L. Cell autonomous sanctions in legumes target ineffective rhizobia in nodules with mixed infections. *Am. J. Bot.* **2017**, *104*, 1299–1312. [CrossRef]
139. Burghardt, L.T. Evolving together, evolving apart: Measuring the fitness of rhizobial bacteria in and out of symbiosis with leguminous plants. *New Phytol.* **2019**. [CrossRef]
140. Denamur, E.; Matic, I. Evolution of mutation rates in bacteria. *Mol. Microbiol.* **2006**, *60*, 820–827. [CrossRef]
141. Raynes, Y.; Sniegowski, P.D. Experimental evolution and the dynamics of genomic mutation rate modifiers. *Heredity (Edinb.)* **2014**, *113*, 375–380. [CrossRef]
142. Foster, P.L. Stress-induced mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol.* **2007**, *42*, 373–397. [CrossRef]
143. Galhardo, R.S.; Hastings, P.J.; Rosenberg, S.M. Mutation as a stress response and the regulation of evolvability. *Crit. Rev. Biochem. Mol. Biol.* **2007**, *42*, 399–435. [CrossRef]
144. Remigi, P.; Capela, D.; Clerissi, C.; Tasse, L.; Torchet, R.; Bouchez, O.; Batut, J.; Cruveiller, S.; Rocha, E.P.; Masson-Boivin, C. Transient hypermutagenesis accelerates the evolution of legume endosymbionts following horizontal gene transfer. *PLoS Biol.* **2014**, *12*, e1001942. [CrossRef] [PubMed]
145. Didelot, X.; Walker, A.S.; Peto, T.E.; Crook, D.W.; Wilson, D.J. Within-host evolution of bacterial pathogens. *Nat. Rev. Microbiol.* **2016**, *14*, 150–162. [CrossRef] [PubMed]

146. Linz, B.; Windsor, H.M.; McGraw, J.J.; Hansen, L.M.; Gajewski, J.P.; Tomsho, L.P.; Hake, C.M.; Solnick, J.V.; Schuster, S.C.; Marshall, B.J. a mutation burst during the acute phase of *Helicobacter pylori* infection in humans and rhesus macaques. *Nat. Commun.* **2014**, *5*, 4165. [CrossRef]
147. Barros-Carvalho, G.A.; Hungria, M.; Lopes, F.M.; Van Sluys, M.A. Brazilian-adapted soybean *Bradyrhizobium* strains uncover IS elements with potential impact on biological nitrogen fixation. *FEMS Microbiol. Lett.* **2019**, *366*, fnz046. [CrossRef] [PubMed]
148. Dupuy, P.; Sauviac, L.; Bruand, C. Stress-inducible NHEJ in bacteria: Function in DNA repair and acquisition of heterologous DNA. *Nucleic Acids Res.* **2019**, *47*, 1335–1349. [CrossRef] [PubMed]
149. Wielgoss, S.; Barrick, J.E.; Tenaillon, O.; Wiser, M.J.; Dittmar, W.J.; Cruveiller, S.; Chane-Woon-Ming, B.; Medigue, C.; Lenski, R.E.; Schneider, D. Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 222–227. [CrossRef] [PubMed]
150. McDonald, M.J.; Hsieh, Y.Y.; Yu, Y.H.; Chang, S.L.; Leu, J.Y. The evolution of low mutation rates in experimental mutator populations of *Saccharomyces cerevisiae*. *Curr. Biol.* **2012**, *22*, 1235–1240. [CrossRef] [PubMed]
151. Swings, T.; Van den Bergh, B.; Wuyts, S.; Oeyen, E.; Voordeckers, K.; Verstrepen, K.J.; Fauvart, M.; Verstraeten, N.; Michiels, J. Adaptive tuning of mutation rates allows fast response to lethal stress in *Escherichia coli*. *Elife* **2017**, *6*, e22939. [CrossRef]
152. Remigi, P.; Masson-Boivin, C.; Rocha, E.P.C. Experimental evolution as a tool to investigate natural processes and molecular functions. *Trends Microbiol.* **2019**, *27*, 623–634. [CrossRef]
153. Yi, X. Experimental evolution and proximate mechanisms in biology. *Synth. Syst. Biotechnol.* **2017**, *2*, 253–258. [CrossRef] [PubMed]
154. Peeters, N.; Guidot, A.; Vailleau, F.; Valls, M. *Ralstonia solanacearum*, a widespread bacterial plant pathogen in the post-genomic era. *Mol. Plant Pathol.* **2013**, *14*, 651–662. [CrossRef] [PubMed]
155. Ray, S.K.; Kumar, R.; Peeters, N.; Boucher, C.; Genin, S. *rpoN1*, but not *rpoN2*, is required for twitching motility, natural competence, growth on nitrate, and virulence of *Ralstonia solanacearum*. *Front. Microbiol.* **2015**, *6*, 229. [CrossRef] [PubMed]
156. Hikichi, Y.; Mori, Y.; Ishikawa, S.; Hayashi, K.; Ohnishi, K.; Kiba, A.; Kai, K. Regulation involved in colonization of intercellular spaces of host plants. *Front. Plant Sci.* **2017**, *8*, 967. [CrossRef] [PubMed]
157. Zhang, Y.; Zhang, W.; Han, L.; Li, J.; Shi, X.; Hikichi, Y.; Ohnishi, K. Involvement of a PadR regulator PrhP on virulence of *Ralstonia solanacearum* by controlling detoxification of phenolic acids and type III secretion system. *Mol. Plant Pathol.* **2019**, *20*, 1477–1490. [CrossRef] [PubMed]
158. Zhang, Y.; Li, J.; Zhang, W.; Shi, H.; Luo, F.; Hikichi, Y.; Shi, X.; Ohnishi, K. a putative LysR-type transcriptional regulator PrhO positively regulates the type III secretion system and contributes to the virulence of *Ralstonia solanacearum*. *Mol. Plant Pathol.* **2018**, *19*, 1808–1819. [CrossRef]
159. Schell, M.A. Control of virulence and pathogenicity genes of *Ralstonia solanacearum* by an elaborate sensory network. *Ann. Rev. Phytopathol.* **2000**, *38*, 263–292. [CrossRef]
160. Clough, S.J.; Lee, K.E.; Schell, M.A.; Denny, T.P. a two-component system in *Ralstonia* (*Pseudomonas*) *solanacearum* modulates production of PhcA-regulated virulence factors in response to 3-hydroxypalmitic acid methyl ester. *J. Bacteriol.* **1997**, *179*, 3639–3648. [CrossRef]
161. Guidot, A.; Jiang, W.; Ferdy, J.-B.; Thebaud, C.; Barberis, P.; Gouzy, J.; Genin, S. Multihost experimental evolution of the pathogen *Ralstonia solanacearum* unveils genes involved in adaptation to plants. *Mol. Biol. Evol.* **2014**, *31*, 2913–2928. [CrossRef]
162. D'Haeze, W.; Holsters, M. Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology* **2002**, *12*, 79R–105R. [CrossRef]
163. Oldroyd, G.E.D.; Murray, J.D.; Poole, P.S.; Downie, J.A. The rules of engagement in the legume-rhizobial symbiosis. *Ann. Rev. Genet.* **2011**, *45*, 119–144. [CrossRef] [PubMed]
164. Daubech, B. Évolution Expérimentale d'un Symbiote de Légumineuse: Étude des Facteurs Génétiques et des Forces de Sélection qui Favorisent ou non l'évolution du Mutualisme. Ph.D. Thesis, Université Toulouse 3 Paul Sabatier, Toulouse, France, 2019.
165. Daubech, B.; Poinsot, V.; Klonowska, A.; Capela, D.; Chaintreuil, C.; Moulin, L.; Marchetti, M.; Masson-Boivin, C. a new nodulation gene involved in the biosynthesis of Nod Factors with an open-chain oxidized terminal residue and in the symbiosis with *Mimosa pudica*. *Mol. Plant Microbe Interact.* **2019**, *32*, 1635–1648. [CrossRef] [PubMed]

166. Sugawara, M.; Takahashi, S.; Umehara, Y.; Iwano, H.; Tsurumaru, H.; Odake, H.; Suzuki, Y.; Kondo, H.; Konno, Y.; Yamakawa, T.; et al. Variation in bradyrhizobial NopP effector determines symbiotic incompatibility with Rj2-soybeans via effector-triggered immunity. *Nat. Commun.* **2018**, *9*, 3139. [CrossRef] [PubMed]
167. Crook, M.B.; Lindsay, D.P.; Biggs, M.B.; Bentley, J.S.; Price, J.C.; Clement, S.C.; Clement, M.J.; Long, S.R.; Griffitts, J.S. Rhizobial plasmids that cause impaired symbiotic nitrogen fixation and enhanced host invasion. *Mol. Plant Microbe Interact.* **2012**, *25*, 1026–1033. [CrossRef]
168. Burghardt, L.T.; Epstein, B.; Guhlin, J.; Nelson, M.S.; Taylor, M.R.; Young, N.D.; Sadowsky, M.J.; Tiffin, P. Select and resequence reveals relative fitness of bacteria in symbiotic and free-living environments. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2425–2430. [CrossRef] [PubMed]
169. Burghardt, L.T.; Trujillo, D.I.; Epstein, B.; Tiffin, P.; Young, N.D. a select and resequence approach reveals strain-specific effects of *Medicago* nodule-specific PLAT-domain genes. *Plant Physiol.* **2020**, *182*, 463–471. [CrossRef]
170. Rogers, C.; Oldroyd, G.E.D. Synthetic biology approaches to engineering the nitrogen symbiosis in cereals. *J. Exp. Bot.* **2014**, *65*, 1939–1946. [CrossRef]
171. Pankievicz, V.C.S.; Irving, T.B.; Maia, L.G.S.; Ané, J.M. Are we there yet? The long walk towards the development of efficient symbiotic associations between nitrogen-fixing bacteria and non-leguminous crops. *BMC Biol.* **2019**, *17*, 99. [CrossRef]