HAL open science

Rhizobium diversity in the light of evolution

Mingxing Tang, Delphine Capela

▶ To cite this version:

Mingxing Tang, Delphine Capela. Rhizobium diversity in the light of evolution. Regulation of Nitrogen-Fixing Symbioses in Legumes, 94, pp.251-288, 2020, Advances in Botanical Research, ⟨10.1016/bs.abr.2019.09.006⟩. ⟨hal-03037688⟩

HAL Id: hal-03037688

https://hal.science/hal-03037688v1

Submitted on 3 Dec 2020

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

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

Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Running title: Rhizobium diversity and evolution

Rhizobium diversity in the light of evolution

Mingxing Tang[1] and Delphine Capela[1*]

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

[*]Corresponding author, e-mail: delphine.capela@inra.fr

Abstract

Rhizobia, nitrogen-fixing symbionts of legumes, are phylogenetically diverse soil bacteria. These bacteria do not share a unique genetic strategy to achieve symbiosis. Instead, they have evolved various mechanisms to induce and invade the root/stem organs where nitrogen fixation occurs. Most rhizobial lineages emerged from horizontal transfer of few essential symbiotic genes located on mobile genetic elements. Besides, rhizobia have extensively recruited their own native functions to optimize symbiosis and refine interaction with host plants. This has likely occurred following symbiotic gene transfer and via genomic adjustments under plant selection pressure. This two-step evolutionary scenario was recently validated by an evolution experiment that progressively turned a plant pathogen into legume symbionts. Acquisition and improvement of the first symbiotic stages, nodulation and infection, were most often gained via the rewiring of the virulence and/or metabolic regulatory network of the ancestral strain.

Keywords: Rhizobium, symbiosis, diversity, adaptation, HGT (horizontal gene transfer), evolution.

22 **Introduction**

23 The generic term rhizobium designates bacteria that share the capacity to establish a nitrogen-fixing
24 symbiosis with legumes. Rhizobia induce, usually on the roots and occasionally on the stems of host
25 plants, the formation of nodules they massively and intracellularly colonize. Several modes of entry of
26 bacteria in plant tissues have been described depending on the symbiotic partner (Madsen et al., 2010).
27 In temperate legumes such as *Medicago* or *Lotus*, rhizobia mainly invade the root at the tip of root hairs,
28 which curl to entrap one or two bacteria and invaginate to form tubular structures called infection
29 threads (ITs). Bacteria grow and divide inside ITs that elongate and ramify to cross the root cell layers
30 towards cortical cells (Gage, 2002). Coordinately with ITs progression, root cortical cells at the basis of
31 the infected root hair divide to form a nodule primordium. Then bacteria exit ITs to colonize the
32 cytoplasm of nodule cells. In 25% of legumes, mainly tropical and subtropical legumes such as *Sesbania*
33 or *Arachis*, bacteria enter the roots intercellularly by "crack entry" at the junctions of main and lateral
34 roots or wounds. Intercellular bacteria trigger plant cell death and form infection pockets, from which
35 bacteria spread either through the formation of ITs and release of bacteria in cortical cells or through a
36 direct invasion of cortical cells (Capoen et al., 2010; Madsen et al., 2010). Another simpler mode of
37 infection was recently described in *Aeschynomene* species where bacteria enter intercellularly without
38 ITs formation nor distant induction of nodule primordium (Bonaldi et al., 2011). In these cases, bacteria
39 invade the roots by a mechanism that resembles the crack entry mechanism and then one bacterium is
40 delivered into a single founder cortical cell, which subsequently divides to form a nodule primordium. In
41 certain Aeschynomene species, in addition to the infection of the core nodule, superficial tissues are
42 invaded by undifferentiated bacteria enclosed in tubular structures that may constitute a reservoir of
43 viable bacteria able to infect new plants after the nodule senescence. Whatever the infection mode is
44 used, all rhizobium-legume associations lead to the formation of symbiosomes. Symbiosomes are
45 organelle-like structures in the cytosol of nodule cells where bacteria surrounded by a plant membrane

46 differentiate into nitrogen-fixing bacteroids. Reduced dinitrogen into ammonia is directly exported and
47 assimilated by the plant that in turn provides bacteria with carbon sources, leading to mutualistic
48 interactions.
49 Rhizobia are very diverse bacteria that have been identified in distant phylogenetically branches,
50 interspersed with non rhizobial bacterium species (Remigi et al., 2016). So far, up to 18 different
51 rhizobium genera and hundreds of species have been described among α- and β-proteobacteria, thus
52 now referred as α- and β-rhizobia (Table 1). The rhizobium genetic diversity is reflected by the diversity
53 of molecular mechanisms these bacteria use during the symbiotic process. Comparative genomic studies
54 of rhizobium symbiotic genes indeed do not support a unique shared genetic strategy for symbiosis in
55 rhizobia (Amadou et al., 2008; Tian et al., 2012). Instead various functions, whose phylogenetic
56 distribution ranges from ubiquitous to lineage- or even strain-specific, have been recruited for symbiosis.
57 The ability of rhizobia to induce a mutualistic symbiosis relies on their capacity to i) trigger a plant
58 developmental stage leading to nodule organogenesis and infection, ii) suppress and/or avoid plant
59 immunity to allow the massive and persistent invasion of bacteria and iii) fix and export nitrogen in
60 exchange to photosynthetates. In this review, we examined the different genetic strategies rhizobia have
61 co-opted to achieve these steps. We explained this diversity by the evolutionary history of these
62 bacteria, which mainly emerged from horizontal gene transfer (HGT) of key symbiotic genes followed by
63 genome remodeling to activate and optimize the symbiotic potential. Finally, we reported a unique
64 evolution experiment replaying the emergence of a new rhizobium under laboratory conditions and
65 validating the proposed rhizobium evolutionary scenario.
66

67 **1. Rhizobia achieve symbiosis via genetically diverse strategies**

68 **1.1. Nodulation: a widespread, yet not unique, Nod Factor strategy**

69 The nodule developmental stage (organogenesis and root infection) is triggered via the activation of an
70 ancestral plant symbiotic signaling pathway that was initially used in most land plants to establish
71 ancient symbioses with mycorrhizal fungi. Notably, this pathway is also recruited in actinorhizal plants to
72 form nitrogen-fixing nodules with *Frankia* actinobacteria (Gherbi et al., 2008; Hocher et al., 2011;
73 Svistoonoff et al., 2013). This common symbiosis signaling pathway (CSSP) is well characterized in the
74 model legumes *Medicago truncatula* (Mt) and *Lotus japonicus* (Lj). It involves a first set of signal
75 transduction proteins MtDMI2/LjSYMRK, and MtDMI1/LjCastor-LjPollux-LjNENA-LjNup85-LjNup133 that
76 produces Ca^{2+} spiking signals following the perception of the microsymbiont, and a second complex
77 composed of MtIPD3/LjCYCLOPS and MtDMI3/LjCCaMK that decodes the calcium signals. In turn, this
78 second complex activates specific downstream transcription factors to induce either mycorrhizal
79 symbiosis or root nodule symbiosis (Gough and Cullimore, 2011; Oldroyd, 2013).
80 Most rhizobia activate the CSSP pathway via the synthesis and secretion of Nod Factors (NFs) (D'Haeze
81 and Holsters, 2002). Rhizobial NFs are the products of enzymes encoded by the nodulation (*nod*) genes
82 (Figure 1A). The expression of *nod* genes is induced by flavonoids, aromatic compounds exuded by host
83 roots, and controlled by NodD transcription regulators that recognize conserved motifs (called *nod*-
84 boxes) in their promoter region. Flavonoids attract bacteria near the roots and bind to the bacterial
85 NodD proteins. In the case of compatible NodD-flavonoid interaction, this leads to DNA bending and
86 further activation of the downstream genes. NodD regulators thus act both as sensors of the plant signal
87 and as transcription regulators of *nod* loci (Schlaman et al., 1992). The NF backbone consists of an
88 oligomer of chitin acylated at the non-reducing terminal residue. This backbone is synthesized by the
89 common NodABC proteins present in all NF-producing rhizobia. In addition, rhizobial strains possess
90 specific pools of *nod*, *nol* or *noe* genes encoding enzymes involved in the attachment of various
91 substituents (Table 2) such as methyl (*nodS*), carbamoyl (*nolO* and *nodU*), acetyl (*nodL, noeT*), sulfate
92 (*nodH*, *nodPQ*), fucose (*nodZ*), O-methyl-fucose (*noeI*), O-sulfate-fucose (*noeE*), O-acetyl-fucose (*nolL*) or

93 arabinose (*noeC*) at different locations of the basal structure (Downie, 2010; Österman et al., 2014). NFs
94 also vary in the length of the backbone and in the nature of the fatty acyl group (D'Haeze and Holsters,
95 2002). Usually rhizobia do not produce only one type of NFs but a set of NFs whose diversity determine
96 the host-range of the strain (D'Haeze and Holsters, 2002; Perret et al., 2000). NFs are exported via a type
97 I secretion system encoded by the *nodIJ* genes, and recognized by a cognate host receptor complex,
98 comprising MtNFP and MtLYK3 in *M. truncatula* (Mt) and LjNFR5 and LjNFR1 in *L. japonicus* (Lj). The
99 perception of NFs initiates the phosphorylation of these receptors which induces the activation of the
100 CSSP (Kouchi et al., 2010; Oldroyd, 2013). Then the CSSP activates the specific transcription factors NSP1,
101 NSP2, NIN, ERN1 and ERN2 coordinating the nodule development and rhizobial infection in the different
102 root tissues (Cerri et al., 2016; Cerri et al., 2017; Vernié et al., 2015).
103 A NF-independent strategy was uncovered a decade ago in photosynthetic *Bradyrhizobium* lacking the
104 *nodABC* genes and nodulating certain *Aeschynomene* species (Chaintreuil et al., 2018; Giraud et al., 2007;
105 Okazaki et al., 2016). The bacterial factors involved in this alternative nodulation mechanism have not
106 been elucidated yet. However two key proteins of the CSSP, SymRK and CCaMK, as well as the LHK1
107 cytokinin receptor required for *nod*-dependent nodulation process were shown to be also required for
108 the *nod*-independent process (Fabre et al., 2015). Recently, a transcriptomic analysis of the *nod*-
109 independent *A. evenia* - *Bradyrhizobium* ORS278 symbiotic interaction identified transcripts of an
110 NFP/NFR5 ortholog and orthologs of most of the genes belonging to the CSSP (MtDMI2/LjSYMRK,
111 LjCastor-MtDMI1/LjPollux-LjNENA-LjNup85-LjNup133, MtIPD3/LjCYCLOPS and MtDMI3/LjCCaMK),
112 supporting the involvement of the same plant signaling pathway in both *nod*-dependent and *nod*-
113 independent symbioses (Gully et al., 2018). It is not clear for the moment whether this NF-independent
114 strategy, so far restricted to some *Aeschynomene Bradyrhizobia*, is either a vestige of an ancient
115 primitive mechanism or a more recently evolved mechanism (Okubo et al., 2012).
116

117 **1.2. Infection: the requirement of surface polysaccharides**

118 In addition to NFs, the progression of the symbiotic interaction requires the synthesis of surface
119 polysaccharides that are essential for both early and late infection stages in all rhizobium-legume
120 symbiotic interactions (Figure 1A). Four main types of rhizobial surface polysaccharides have been shown
121 to play a role in symbiosis: secreted exopolysaccharides (EPS), capsular polysaccharides (KPS) and
122 lipopolysaccharides (LPS), both anchored in the outer membrane, and cyclic β-glucans mostly located in
123 the periplasm (Downie, 2010; Fraysse et al., 2003; López-Baena et al., 2016). The symbiotic role of these
124 molecules often depends on the rhizobium species and plant host genotype.

125 Although their precise functions have not been resolved, it was proposed that surface polysaccharides
126 protect rhizobia against a variety of biotic and abiotic stresses encountered under symbiotic conditions
127 (osmotic stress, acidification, antimicrobial peptides, or reactive oxygen species) (Arnold et al., 2018;
128 D'Haeze and Holsters, 2004; Davies and Walker, 2007; Geddes et al., 2014; Ingram-Smith and Miller,
129 1998; Lehman and Long, 2013; Miller-Williams et al., 2006). These molecules likely ensure the protection
130 of rhizobia either by forming a physical diffusion barrier around bacteria or by actively inducing pathways
131 that prevent plant defense responses, a prerequisite for nitrogen fixation (D'Antuono et al., 2008;
132 D'Haeze and Holsters, 2004; Jaszek et al., 2014; Jones et al., 2008). The active signaling function of EPS
133 was recently provided by the identification of an EPS receptor in *L. japonicus* (Kawaharada et al., 2015;
134 Kawaharada et al., 2017). Future studies should reveal whether other plant receptors not only to EPS but
135 also to the other rhizobial polysaccharides, such as KPS and LPS, also contribute to the symbiosis
136 signaling pathway.

137

138 *Exopolysaccharides (EPS)*

139 Rhizobial exopolysaccharides play a major role in the initiation and elongation of infection threads (ITs).
140 In *S. meliloti*, two distinct EPSs have been reported, the EPS I or succinoglycan and EPS II or

141 galactoglucan. The low molecular weight form of EPS I was shown to play a major role in the invasion of
142 *Medicago* nodules while EPS II have a more moderate role (Battisti et al., 1992; Cheng and Walker,
143 1998). The *exo* genes required for the synthesis and export of EPS I are well conserved in *Sinorhizobium*,
144 *Rhizobium*, and *Mesorhizobium* while most of them are absent in *Bradyrhizobium* and *Azorhizobium*, two
145 more distant genera of rhizobia, as well as in β-rhizobia (Black et al., 2012) (Table 2). However EPS I are
146 not always essential for infection. For example in *S. fredii*, mutations in *exo* genes do not alter symbiosis
147 (Kim et al., 1989). In this case, other surface polysaccharides, the KPS, are of main importance for
148 infection. Also the *Vicia sativa* (vetch) - *Rhizobium leguminosarum* symbiotic interaction relies on
149 another type of EPS synthesized by the products of *pss* genes specifically present in the *Rhizobium* genus
150 (Table 2) (Black et al., 2012; van Workum et al., 1998).
151 A large diversity in EPS chemical structures can be found among rhizobia (van Workum et al., 1998;
152 Wang et al., 2018). They consist in repeating unit of heteropolysaccharides which varies in the
153 composition of sugars, their linkage within the unit, the degree of polymerization and the presence of
154 non-carbohydrate substituents such as pyruvate, succinate or acetate (Skorupska et al., 2006). A number
155 of evidences indicated that the chemical structure of EPS influences infection efficiency in a host-
156 dependent manner, suggesting an active role of EPS as signal molecules in the rhizobium-legume
157 interactions. For example, an *exoU* mutant in *M. loti* R7A is unable to induce infected nodules on the Gifu
158 ecotype of *L. japonicus,* while this mutant displayed only reduced infection efficiency on other ecotypes
159 of *L. japonicus* such as MG-20 (Kelly et al., 2013). In the same line, complementation of EPS deficient
160 strains by natural heterologous EPS produced by other rhizobia does not always lead to a successful
161 infection. For examples, EPS from *R. tropici* or *R. etli* restore the infection ability of *Rhizobium*
162 *leguminosarum* bv. *viciae pssD* mutants but EPS from *S. meliloti*, *Sinorhizobium fredii* NGR234 or *B.*
163 *diazoefficiens* cannot (van Workum et al., 1998). In the same way, EPS from *Rhizobium leguminosarum*
164 bv. *trifolii* and *S. fredii* NGR234 cannot restore the defective infection phenotype of *S. meliloti exo*

165 mutants on alfalfa (Battisti et al., 1992). The recent identification in *L. japonicus* of EPR3, a specific EPS
166 receptor, that mediates the progression of ITs through the root cortex and nodule primordium
167 confirmed the role of EPS as important signaling molecules (Kawaharada et al., 2015; Kawaharada et al.,
168 2017; Muszyński et al., 2016). This receptor is part of the symbiotic signal transduction pathway since its
169 expression is activated downstream from the CSSP (Kawaharada et al., 2017). Interestingly, EPR3
170 orthologs were found in other legumes, such as *M. truncatula,* soybean as well as *Arachis hypogaea*
171 whose infection is mediated by crack entry but not by ITs (Karmakar et al., 2019; Kawaharada et al.,
172 2015).
173
174 *Capsular polysaccharides (KPS)*
175 Capsular polysaccharides are composed of dimeric repeating unit containing one hexose linked with a
176 Kdo (3-deoxy-D-manno-oct-2-ulosonic acid). The number of *kps* gene clusters is variable among
177 rhizobium genomes even within the same rhizobium species (Kiss et al., 2001; Sugawara et al., 2013).
178 Their importance in symbiosis is also variable. For instance, in *S. fredii* HH103*,* KPS play a major role in
179 nodule formation and growth (Parada et al., 2006) whereas in *S. meliloti* Rm1021, they have only a minor
180 effect (Pellock et al., 2000).
181
182 *Lipopolysaccharides (LPS)*
183 LPS are major constituents of the outer membrane in gram-negative bacteria. They are composed of
184 three parts, a core polysaccharide linked to an O-chain polysaccharide (or O-antigen) and a lipid A
185 (Serrato, 2014). Mutants of rhizobia deficient in LPS biosynthesis are affected either in early infection
186 stages with the formation of abortive ITs or in late infection stages with a rapid lysis of bacteria after
187 their release in the cytoplasm of nodule cells both leading to the formation of non fixing nodules
188 (Broughton et al., 2006; Campbell et al., 2002; Noel et al., 2000). In gram-negative bacteria, the O-

189 antigen is the main immunogenic feature and the most variable part of LPS. Mutants altering the
190 composition of O-antigen in *R. etli* and *S. fredii* NGR234 are impaired in IT development and nitrogen
191 fixation (Reuhs et al., 2005). Interestingly, in these two rhizobia, modifications in the O-chain of LPS
192 occurred under symbiotic conditions (Broughton et al., 2006; Kobayashi et al., 2004; Noel et al., 2004;
193 Ojeda et al., 2010).
194
195 *Cyclic β-glucans*
196 All bacteria from the Rhizobiales order are able to synthesize cyclic β-glucans (CGs). CGs are neutral
197 cyclic homopolymers of β-linked glucose residues often substituted by phosphoglycerol, phosphocholin
198 or succinyl groups. The linkage between the glucose residues, either β(1,2) or β(1,3) or β(1,6), is variable
199 among legume symbionts, however, there is no correlation between the nature of the linkage and their
200 role on host plants (Breedveld and Miller, 1994). For instance, different soybean symbionts were shown
201 to produce different CGs (Crespo-Rivas et al., 2009; Rolin et al., 1992). In α-rhizobia, CGs are essential for
202 the invasion of nodules. Mutants in the synthesis or export in the periplasm of CGs elicit small, white
203 pseudonodules containing abortive ITs and that fail to fix nitrogen (Bhagwat et al., 1996; Bhagwat et al.,
204 1992; Dylan et al., 1990).
205
206 **1.3. Nitrogen fixation: a similar set of *nif-fix* genes for all rhizobia**
207 The reduction of dinitrogen into ammonia is achieved thanks to the nitrogenase enzymatic complex. All
208 rhizobia possess the same type of nitrogenase, the molybdenum-dependent nitrogenase, encoded by
209 the *nif* genes. The number of *nif* genes found in rhizobia is lower than that of free-living diazotrophs,
210 such as *K. pneumoniae*, which possesses 18 structural *nif* genes. A minimal set of genes essential for
211 nitrogen fixation *nifHDK* (nitrogenase enzyme complex) and *nifENB* (FeMoco biosynthesis) is present in
212 all rhizobia. Some other *nif* genes, *nifTXQWZSUV*, involved in the FeMoco synthesis, the Fe-S cluster

213 synthesis and electron transport have been found in rhizobium genomes but their presence is variable
214 (Table 2). Missing *nif* genes compared to free-living diazotrophs might be replaced by housekeeping
215 paralogs (*iscA*/*nifU* and *iscS*/*nifS*) or encoded by yet unidentified genes or, like *nifV*, can be
216 complemented by the FEN1 homocitrate synthase gene of host plants (Hakoyama et al., 2009; Nouwen
217 et al., 2017). Like in free-living diazotrophs, the expression of *nif* genes in rhizobia is directly controlled
218 by the transcription regulator NifA. Both *nifA* transcription and NifA activity are only allowed under very
219 low oxygen concentration (Bobik et al., 2006; Sciotti et al., 2003). This double control of NifA by oxygen
220 may prevent the synthesis of nitrogenase in conditions where this enzyme would be inactive. Under
221 symbiotic conditions, rhizobia have to cope with the conflicting situation where the nitrogenase is
222 extremely sensitive to oxygen while bacteria are strictly aerobic. Moreover, the ATP-dependent and high
223 energy-consuming process of nitrogen fixation necessitates rapid respiration rates of bacteroids. This
224 problem is solved by the plant that provides bacteria with carbon sources and produces high amounts of
225 leghemoglobins, plant proteins transporting and delivering oxygen to bacteroids while maintaining the
226 concentration of free oxygen at very low levels. Under these microaerobic conditions, the *ccoNOQP*
227 (formerly *fixNOQP*) and *ccoGHIS* (formerly *fixGHIS*) genes encoding a cytochrome *cbb*3 oxidase with high
228 affinity for oxygen are essential for nitrogen fixation. In some rhizobia able to fix nitrogen under both
229 free-living and symbiotic conditions, such as *A. caulinodans* and the two beta-rhizobia *P. phymatum* and
230 *P. tuberum*, the cytochrome *bd* terminal oxidase is also or alternatively used (De Meyer et al., 2016).
231 Last, the *fixABCX* genes encoding a specific delivery system of electrons to nitrogenase (Ledbetter et al.,
232 2017) are conserved in most rhizobia (Black et al., 2012).
233
234 **1.4. Optimization of symbiosis through a diversity of molecular and regulatory mechanisms**
235 *Specific symbiotic functions*

236 Besides the above-mentioned essential genetic determinants for symbiosis shared by most rhizobia,
237 additional rhizobial factors allow or optimize symbiosis depending on the rhizobium strain and host
238 genotype (Figure 1).
239 Proteins secretion systems such as T3SS, T4SS and T6SS are examples of accessory functions that
240 modulate symbiosis (Bladergroen et al., 2003; Deakin and Broughton, 2009; Hubber et al., 2004). Type III
241 secretion systems (T3SS), found in many rhizobium strains belonging to *Sinorhizobium*, *Mesorhizobium*,
242 *Bradyrhizobium*, and *Cupriavidus* genera (Amadou et al., 2008; Hubber et al., 2004; Krause et al., 2002;
243 Okazaki et al., 2016; Okazaki et al., 2009; Perret et al., 2000; Sánchez et al., 2009; Viprey et al., 1998),
244 were shown to promote nodulation in several rhizobium-legume interactions. For example, inactivation
245 of the T3SS in the wide host-range rhizobium *S. fredii* NGR234 decreases its nodulation efficiency on
246 many host plants like *Tephrosia vogelii*, *Lablab purpureus* and *Flemingia congest* (Marie et al., 2003;
247 Viprey et al., 1998). Likewise, mutations in the *M. loti* MAFF303099 T3SS affects its competitiveness on
248 *Lotus* (Sánchez et al., 2009). Rhizobial T3SSs resemble those found in pathogenic bacteria and deliver
249 effector proteins, called Nops (Nodule outer protein), directly into the host cells. The set of effector
250 proteins largely vary between rhizobium species and strains. Few Nops have been characterized in
251 details but some of them were clearly shown to interfere with plant defense responses. Expression of
252 the NGR234 effector NopL in the non-host tobacco reduces the expression of pathogen-related proteins
253 by interfering with the MAPK (mitogen-activated protein kinase) signaling while expression of NopM, an
254 E3 ubiquitin ligase domain effector, reduces the production of reactive oxygen species in tobacco
255 (Bartsev et al., 2004; Xin et al., 2012; Zhang et al., 2011). However these secretion systems are not
256 always essential for symbiosis and may even be incompatible with some hosts (Nelson and Sadowsky,
257 2015; Saad et al., 2012; Viprey et al., 1998; Zhao et al., 2018). In these latter cases, it is thought that
258 plant R genes recognize bacterial effectors which induces plant defenses and prevents nodulation, like

259 effector-triggered immune responses described in plant-pathogen interactions (Fan et al., 2017; Nelson
260 and Sadowsky, 2015; Sugawara et al., 2018; Tang et al., 2016; Yang et al., 2010).
261 In addition to their role in the suppression of plant defense reactions, it was also recently proposed that
262 Nop effectors may participate in the activation of the NF signaling pathway. This hypothesis came with
263 the discovery that the T3SS of *Bradyrhizobium elkanii* USDA61 can bypass the NF-perception to induce
264 nodule formation on soybean, a normally NF-dependent process (Okazaki et al., 2013). Still, the
265 nodulation was poorly efficient in this case and bacterial invasion was achieved through an abnormal
266 intercellular mode of infection of root tissues. Interestingly, the T3SS of some non-photosynthetic
267 *Bradyrhizobium* sp. induces nodule formation on the NF-independent *Aeschynomene* legumes,
268 supporting their putative role in the activation of the CSSP (Okazaki et al., 2016).
269 Another example of bacterial gene specifically required in particular rhizobium-host interactions is
270 *bacA*/*bclA* (Guefrachi et al., 2015; Haag et al., 2011). These genes encode peptide transporters that are
271 essential for symbioses with legumes producing antimicrobial peptides known as NCR (Nodule specific
272 Cysteine-Rich) peptides. NCR-producing legumes were found in two distinct phylogenetic clades, the
273 Inverted Repeat-Lacking Clade (IRLC) and *Aeschynomene* spp. from the Dalbergioid clade (Czernic et al.,
274 2015; Mergaert et al., 2003). Rhizobia interacting with these two clades of legumes, *Sinorhizobium* and
275 *Bradyrhizobium* species, respectively, undergo a terminal differentiation process in nodule cells which is
276 induced by the action of hundreds of NCR peptides and requires the bacterial BacA/BclA peptide uptake
277 system. Indeed mutants of these genes in *Sinorhizobium* and *Bradyrhizobium* do not survive in nodule
278 cells (Van de Velde et al., 2010). Consistently, these genes are completely dispensable in rhizobia
279 interacting with non NCR-producing legumes (Maruya and Saeki, 2010).
280
281 *Fine-tuned regulation of symbiotic functions*

282 A crucial level of optimization of symbiosis occurs through regulatory systems that adjust the expression
283 of symbiotic functions to environmental conditions (Figure 1). A strict and attuned regulation of Nod-
284 factor production, exopolysaccharide biosynthesis and nitrogen fixation genes is required for successful
285 symbiosis. Given the high plasticity of regulatory networks in bacteria, symbiotic regulators are plentiful
286 and noticeably differ among rhizobia.
287 Host-mediated activation of nodulation is fulfilled by NodD proteins, LysR-type transcription regulators,
288 whose activity relies on the binding with compatible compounds exuded by the plants, usually flavonoids
289 or isoflavonoids. Rhizobia often possess multiple NodD proteins that integrate different signals, not only
290 diverse plant inducers but also physiological signals like the nitrogen status of the cell (*e.g*. NodD3 in *S.*
291 *meliloti*) or stress (e.g. NodD3 and NodD5 respond to salt stress in *R. tropici* CIAT899) (del Cerro et al.,
292 2015; Dusha and Kondorosi, 1993; Kelly et al., 2018; Kobayashi et al., 2004). Some NodD proteins also act
293 as repressors of *nod* genes. This is the case of NodD2 in *B. diazoefficiens*, which represses NF production
294 in a quorum-sensing dependent way (Garcia et al., 1996; Loh and Stacey, 2003; Loh et al., 2001), or
295 NodD2 in *S. fredii* NGR234 which down-regulates the *nod* gene activator NodD1 (Fellay et al., 1998). The
296 co-modulation of *nod* gene expression by multiple NodDs regulators likely helps rhizobia to expand their
297 host-range, nodulate under specific (harsh) environmental conditions and maintain a tight regulation of
298 NF synthesis.
299 Biosynthesis of exopolysaccharides is a striking example of function subject to a complex regulation
300 (Janczarek, 2011). In *S. meliloti,* in which EPS regulation was extensively studied, a dozen of regulatory
301 proteins affecting either positively (MucR, SyrM, ExoD, SinI/SinR/ExpR for EPS I and WggR,
302 SinI/SinR/ExpR, PhoB/PhoR for EPS II) or negatively (*e.g*. ExoX, ExsB, CbrA, NtrX, EmrR, ExoR/ExoS/ChvI,
303 EmmABC for EPS I and MucR for EPS II) EPS production has been identified. All these regulators respond
304 to a diversity of physiological conditions. EPS I synthesis is enhanced under ammonium or sulfur
305 starvation, high phosphate concentration, hyperosmotic stress as well as high-cell density and repressed

306 by succinate-mediated catabolite repression, while EPS-II production is promoted under phosphate
307 starvation and high-cell density. Many rhizobial EPS regulatory systems are species-specific (Bonomi et
308 al., 2012 ; Liu et al., 2018 ; Nishihata et al., 2018). The continued discovery of new regulators affecting
309 the production of expolysaccharides in rhizobia accounts for the increasingly complex regulation of this
310 function (Barnett and Long, 2018; Schäper et al., 2017). Most proteins controlling EPS synthesis encode
311 global regulatory systems that not only control EPS synthesis but also affect multiple bacterial behaviors.
312 Strikingly, expression of motility genes are frequently inversely correlated with EPS biosynthesis genes
313 (Barnett and Long, 2018). The three component system ExoR/ExoS/ChvI (RSI) is one crucial regulatory
314 system that switches bacteria from motile free-living cells to non-motile host-invading cells by up and
315 down-regulating the expression of hundreds of genes, including EPS I and flagellum genes (Bélanger et
316 al., 2009). ExoR is a protein secreted to the periplasm, whereas *exoS* and *chvI* encode a histidine kinase
317 sensory protein and a response regulator, respectively (Cheng and Walker, 1998). Orthologs of the full
318 set of RSI genes were uniquely found and conserved in Rhizobiales, an order containing numerous
319 symbionts and pathogenic bacteria. The phylogenetic co-occurrence and conservation of the three
320 components of the system strongly suggests that it is a key adaptive pathway in this order (Heavner et
321 al., 2015). Likewise, the other three-component system EmmABC, that plays an important role in
322 adaptation in *S. meliloti* (Morris and González, 2009), was not found beyond the Rhizobiales (Barnett and
323 Long, 2018).
324 Quorum sensing (QS) systems also affect EPS biosynthesis as well as a wide spectrum of other
325 physiological traits that vary among rhizobial species, including biofilm formation, swarming motility,
326 Type III secretion system, plasmid transfer, cell division, metabolism and transport (Calatrava-Morales et
327 al., 2018; Hoang et al., 2004; Krysciak et al., 2014). Their impact on symbiosis can be either neutral
328 (Krysciak et al., 2014), positive (Gurich and González, 2009; Yang et al., 2009; Zheng et al., 2006) or
329 negative (Rodelas et al., 1999; Rosemeyer et al., 1998). In *S. meliloti* and *R. etli*, mutants in autoinducer

330 synthase genes, *sinI* and *cinI* respectively, were unable to invade nodules efficiently and were affected in
331 symbiosome development and nitrogen fixation (Daniels et al., 2002; Gurich and González, 2009).
332 Nodulation was also severely impaired in QS mutants of *Mesorhizobium loti* and *Mesorhizobium*
333 *tianshanense* on their respective host plants *Lotus corniculatus* and *Glycyrrhiza uralensis* (Yang et al.,
334 2009; Zheng et al., 2006). However there are also examples where mutations in QS systems have no
335 impact on symbiosis. For examples, the TraI/R and NgrI/R QS systems in *S. fredii* NGR234 as well as the
336 BraI/R QS system in the β-rhizobium *P. phymatum* do not affect nodulation on respective host plants,
337 peas, common bean and *M. pudica*, respectively, although these systems control exopolysaccharide
338 production (Coutinho et al., 2013; Krysciak et al., 2014). Conversely, mutants in the autoinducer synthase
339 genes *rhiI* and *raiI*, in *R. leguminosarum* and *R. etli* respectively, induce significantly more nodules than
340 the wild-type strain on the host plants, peas and common bean (Rodelas et al., 1999; Rosemeyer et al.,
341 1998). QS systems may thus be dispensable or not for symbiosis, probably depending on the genes they
342 control and the complexity of the regulatory network to which they belong.
343 Another interesting example of regulatory process that modulates symbiosis in a rhizobium-host
344 dependent manner is the ppGpp-mediated stringent response, which controls diverse cellular functions
345 in response to environmental stresses, especially nutrient starvation (Irving and Corrigan, 2018). The
346 level of intracellular ppGpp depends on the ppGpp synthase and hydrolase activities of enzymes of the
347 RSH superfamily. Recently, a *relA* mutant in *S. meliloti*, that is unable to initiate the stringent response,
348 fails to nodulate the host plant *M. sativa*, but successfully invades *M. truncatula* nodules with, however,
349 a reduced nitrogen fixation capacity (Wippel and Long, 2018). In *Rhizobium etli* CE3, a *relA* mutant led to
350 a strongly reduced capacity in both nodulation and nitrogen fixation (Calderón-Flores et al., 2005), while
351 another strain of *R. etli,* CNPAF512, deficient in ppGpp production, is only impaired at later stage of
352 symbiosis, showing abnormal bacteroid development and significant decrease in nitrogen fixation (Moris
353 et al., 2005).

354 Finally, nitrogen fixation genes are under the control of the master regulator NifA in both α and β-
355 rhizobia. Upstream from NifA, the regulatory cascade differs among rhizobia. At the top of this cascade,
356 the FixLJ two component system, responding to microaerobic conditions, is conserved in α-rhizobia but
357 absent in β-rhizobia (Amadou et al., 2008; Cosseau and Batut, 2004). In *Sinorhizobium meliloti*, it
358 activates the two intermediate regulatory genes *fixK* and *nifA* (Bobik et al., 2006). FixK belongs to the
359 Crp/Fnr family of transcriptional regulators and controls, among other genes, the *ccoNOQP* and *ccoGHIS*
360 operons. In *A. caulinodans*, a rhizobium able to fix nitrogen under both free-living and symbiotic
361 conditions, NifA is under the double control of oxygen and ammonia concentrations through the FixK
362 and NtrBC regulators, respectively (Kaminski and Elmerich, 1998). In *B. diazoefficiens*, *nifA* transcription
363 is largely independent from FixJ and oxygen concentration but instead is activated by the two-
364 component system RegSR which responds to an unknown signal (Bauer et al., 1998). Interestingly, in this
365 rhizobium, many genes not directly related to nitrogen fixation are controlled by NifA (Hauser et al.,
366 2007).
367

368 **2. Parallel evolution as a source of rhizobial diversity**

369 **2.1. The spread of mutualistic traits across diverse bacterial lineages occurred through repeated**
370 **horizontal transfers of symbiotic genes**

371 A large body of evidence indicates that NF-dependent rhizobia emerged once and have diversified from
372 repeated and independent horizontal transfers of symbiosis genes across species and genera.
373 First, the sequencing of numerous rhizobium genomes revealed that essential symbiotic genes, in
374 particular those involved in NFs synthesis (*nod* genes) and nitrogen fixation (*nif-fix* genes), very often, if
375 not always for nodulation genes, reside in clusters on highly mobile genetic elements either genomic
376 islands or plasmids (MacLean et al., 2007).

377 Moreover, the common rhizobial *nodABC* genes are monophyletic (Chen et al., 2003; Gyaneshwar et al.,
378 2011; Laguerre et al., 2001), while rhizobia are polyphyletic, leading to incongruences between
379 rhizobium core genes and symbiosis genes (Andrews et al., 2018). Indeed phylogenetic studies have
380 highlighted the importance of horizontal transfers of symbiosis genes both within and between rhizobial
381 genera.
382 Finally, transfers of symbiotic genomic islands from a rhizobium to non symbiotic soil bacteria were
383 observed in nature. In the pioneer experiment of Sullivan *et al.*, a stand of *Lotus corniculatus* was
384 cultivated in a field devoid of natural rhizobia able to nodulate this legume and inoculated with one *M.*
385 *loti* compatible symbiotic strain. After few years, the original inoculant could not be retrieved from *L.*
386 *corniculatus* nodules. Instead nodules were occupied by local non symbiotic *Mesorhizobium* strains that
387 had all acquired the symbiotic island of the original inoculant (Sullivan and Ronson, 1998; Sullivan et al.,
388 1995). Similar observations were made with *Biserrula pelecinus* symbionts in Australia and soybean
389 symbionts in Brazil (Barcellos et al., 2007; Nandasena et al., 2007). The sequencing of a rhizobium
390 population isolated from the rhizosphere and nodules of beans cultivated in the same area provides
391 another evidence that symbiosis plasmids and islands can spread rapidly within a soil population
392 although these transfers occur preferentially between closely related species (Pérez Carrascal et al.,
393 2016).
394 HGT have thus clearly played a major role in the dissemination of the rhizobial mutualistic traits among
395 α- and β-proteobacteria. Although the origin of *nod* genes is not yet completely elucidated, it was
396 recently proposed that the common canonical *nodABC* genes originate from certain *Frankia* symbionts
397 (Ktari et al., 2017; Persson, 2015). In turn, accessory *nod* genes (involved in NF decorations) could have
398 been recruited via duplications of resident genes encoding housekeeping functions during evolution. In
399 *S. meliloti* strain 1021, the plasmidic *nodP1Q1* genes involved in the sulfation of NFs is thought to have
400 arisen from a transposition of the ancestral genomic copy *nodP2Q2* (paralogs of the *cysDNC* genes)

401 (Capela et al., 2005; Galibert et al., 2001). The NF transport system encoded by the highly conserved
402 *nodIJ* genes is another example of genes that were first duplicated within β-rhizobia and then
403 transferred to α-rhizobia (Aoki et al., 2013).
404
405 **2.2. HGT is likely promoted in the plant environment**
406 In nature, HGT can be mediated by different processes such as conjugation of plasmids, phage
407 transduction, natural transformation, or integrative and conjugative elements (ICEs). Symbiosis genes
408 cluster preferentially in ICEs (chromosomal islands) in rhizobial species of *Mesorhizobium,*
409 *Bradyrhizobium* and *Azorhizobium* genera*,* while large symbiotic plasmids (> 0.2 Mb) have been mainly
410 found in rhizobial species of *Rhizobium*, *Sinorhizobium* (*Ensifer*) as well as the β-rhizobial
411 *Paraburkholderia* and *Cupriavidus* genera.
412 Very low transfer frequencies of ICEs and symbiotic plasmids were observed under laboratory
413 conditions, indicating that these transfers are tightly controlled (Bañuelos-Vazquez et al., 2017; Delavat
414 et al., 2017; Ding and Hynes, 2009). Two main regulatory mechanisms of conjugative transfer were
415 reported for rhizobial symbiotic plasmids in *Rhizobium* and *Sinorhizobium* genera. The first mechanism is
416 mediated by the TraI/TraR/TraM quorum sensing system. In high cell density, quorum sensing molecules
417 produced by the autoinducer synthase TraI binds to the regulator TraR which, in turn, activates the
418 *tra*/*trb*/*virB* (encoding functions necessary for the DNA transfer including the conjugative T4SS) and *rep*
419 (encoding genes necessary for segregation/partition and replication of the plasmid) genes as well as
420 *traM,* a gene encoding an anti-activator of TraR. Additional levels of regulation acting either in parallel or
421 upstream the TraI/TraR/TraM system were reported in several rhizobia. For examples, in *R.*
422 *leguminosarum*, the BisR response regulator and the CinI/CinR quorum sensing system regulate the
423 expression of *traR* (Danino et al., 2003) while in *R. etli,* the LuxR-type regulator CinR positively regulate
424 the expression of *traI* (Tun-Garrido et al., 2003). A second mechanism of regulation, independent of

425 quorum sensing, described in *R. etli* CFN42, involves the RctA repressor of the *tra* and *virB* genes (Pérez-
426 Mendoza et al., 2005) and another regulator RctB that antagonizes the repressive effect of RctA.
427 Additionally, in *S. meliloti*, the expression of *rctB* is repressed by the regulator RctR and activated by the
428 response regulator RctC in response to a yet unidentified signal (Nogales et al., 2013). Homologs of RctR
429 and RctC are present in *S. fredii* NGR234 but are absent in *R. etli*, showing the existence of different
430 signaling cascades and regulation mechanisms of plasmid transfer in rhizobia.
431 The transfer of ICEs involves the excision and circularization of the DNA region which can replicate by a
432 rolling circle mode. A single stranded DNA copy is then transferred to the recipient bacterium using a
433 conjugative T4SS and integrated at specific conserved sites usually near tRNA genes. ICEs encode all the
434 necessary functions for the excision, insertion and conjugation process. In *Mesorhizobium* species,
435 symbiosis ICEs have been found either as a large genomic DNA region of *ca.* 500 kb or as a tripartite
436 element located at three separate chromosomal regions, which reassemble into a single element before
437 the transfer. The complex regulation of ICE transfer was deeply studied in *M. loti* R7A. Activation of
438 excision and transfer of this ICE is mainly orchestrated by the TraI/TraR quorum sensing system which
439 activates FseA, the master regulator of the *rdfS* excisionase gene. This activation is however prevented in
440 most cells through three additional levels of regulation including i) an inhibition of the quorum sensor
441 TraR by the regulator QseM, ii) a low frequency event of ribosomal frameshifting necessary for the
442 translation of the complete FseA protein, and iii) an anti-activation of FseA by QseM preventing
443 activation of *rdfS* (Ramsay et al., 2009; Ramsay et al., 2015). Thus the ICE of *M. loti* R7A transfer only
444 occurs in a population of cells in which both *qseM* expression is repressed and FseA is translated.
445 These tight and complex regulations of DNA transfers contrast with the transfer frequencies observed in
446 natural populations of legume symbionts. It suggests that either environmental factors in the plant
447 rhizosphere promote the transfers of symbiosis genes or rare transfers are rapidly fixed in a population
448 when it provides a strong adaptive advantage, such as the capacity to infect and multiply within legume

449 plants. The role of plant factors in the stimulation of DNA transfer in the rhizosphere was recently
450 provided by the study of Ling *et al*. demonstrating that the HGT of the symbiosis ICE of *A. caulinodans* is
451 triggered by plant flavonoids sensed by a LysR-type transcription regulator, AhaR, controlling the
452 expression of the *intC* integrase gene (Ling et al., 2016). Root exudates were also shown to enhance
453 conjugal plasmid transfers among bacteria in the rhizosphere, indicating that HGTs can be favored in the
454 rhizospheric environment (Mølbak et al., 2007).
455
456 **2.3. Optimization of the symbiosis potential *via* post-transfer adaptations**
457 Horizontal acquisition of symbiosis genes is not always sufficient to immediately convert a recipient
458 bacterium into an efficient legume symbiont and multiple genomic adjustments may be required to
459 achieve symbiosis (Figure 2A). In line with this, HGT observed in fields sometimes results in the evolution
460 of novel rhizobia that are suboptimal for nitrogen fixation. For instance, introduction of the herbaceous
461 legume *Biserrula pelecinus* in Western Australian together with its compatible symbiont *Mesorhizobium*
462 *ciceri* bv. *biserrulae* led to the rapid evolution of novel *Mesorhizobium* strains that have acquired the
463 symbiosis genomic island from *M. ciceri* inoculant but were either ineffective or fix nitrogen poorly. This
464 suggests that the symbiotic potential newly acquired need to be optimized in these strains to produce
465 efficient rhizobia (Nandasena et al., 2007). Also experimental transfers of symbiotic genes to distant non
466 rhizobial bacteria in the lab were usually unsuccessful leading either to non-nodulating strains or to
467 strains able to induce ineffective pseudonodules (Hirsch et al., 1984; Marchetti et al., 2010).
468 One optimization step may rely on the integration of symbiosis genes into the endogenous regulatory
469 circuit. This phenomenon is illustrated by the diversity of regulatory systems controlling nitrogen fixation
470 genes in rhizobia (Dixon and Kahn, 2004). Yet the incorporation of laterally acquired genes into the
471 existing regulatory network was shown to be a long process taking many million years (Lercher and Pál,
472 2008).

473 In addition to the transferred symbiotic genes, rhizobia largely recruit resident genes for symbiosis.
474 Missing genes for nodulation or nitrogen fixation can be recruited from housekeeping genes either
475 directly (*glmS* can replace *nodM, iscA* and *iscS* can replace *nifU* and *nifS,* respectively) or following a
476 duplication and integration into the symbiotic region (*glmS*/*nodM* or *cysDNC*/*nodPQ*) (Galibert et al.,
477 2001; Schwedock and Long, 1992). As mentioned in the previous section, protein secretion systems,
478 although not essential for symbiosis, were frequently recruited in rhizobia to assist NFs to activate the
479 CSSP and/or reduce plant defense responses. Integration of these systems in the nodulation regulatory
480 circuit to coordinate their expression with *nod* genes confirms their important role in the symbiotic
481 adaptation (Hubber et al., 2007; Kobayashi et al., 2004; Nelson et al., 2017). In addition, several
482 examples of ubiquitous functions required for nitrogen fixation, like heme biosynthesis or central
483 metabolism functions, are encoded by native genes that are used to fulfill symbiosis. The phylogenetic
484 analysis of *ca*. 500 genes known to be involved in symbiosis collected from 41 rhizobial strains revealed
485 extensive recruitment of lineage-specific genes (Tian et al., 2012). Striking examples of utilization of local
486 genes for symbiosis are the *exo*, *exs*, *lps*, or *pss* genes, essential for infection and whose phylogeny is
487 consistent with the phylogeny of species (Black et al., 2012; Wang et al., 2018). Some rhizobia have also
488 exploited their intrinsic metabolic properties, such as methylotrophy in *Methylobacterium nodulans* or
489 photosynthesis in *Bradyrhizobium* sp. ORS278, to serve symbiosis (Giraud et al., 2000; Jourand et al.,
490 2005).
491 Other post-transfer adaptive evolution may occur through diverse processes such as allelic variation,
492 gene inactivation or deletion, or additional horizontal gene acquisition. The extent of genomic
493 modifications required to adapt to legume symbiosis is not known and probably depends on the
494 recipient genome and the host genotype. We can hypothesize that symbiosis gene transfers between
495 close and pre-adapted genomes could be immediately efficient while transfers between distant genera
496 would necessitate further adaptation steps to enable symbiosis. Long-term co-evolution of rhizobia and

497 their hosts should also increase the number of genome adjustments to refine the mutualistic
498 interactions.
499

500 ## 3. Insights into the emergence of rhizobia through experimental evolution

501 ### 3.1. An experimental design reproducing the natural evolutionary scenario of rhizobia

502 Some years ago, an evolution experiment was designed to mimic the two-step natural evolutionary
503 scenario of rhizobia and replay the emergence of a new rhizobial genus under laboratory conditions. This
504 experiment consisted in the transfer of essential symbiotic genes from a legume symbiont into a non
505 rhizobium strain followed by further evolution of the resulting strain under the plant selection pressure.
506 The rhizobium *Cupriavidus taiwanensis* and the plant pathogen *Ralstonia solanacearum*, two
507 phylogenetically relatively distant bacteria, were chosen as donor and recipient of symbiosis genes,
508 respectively. The symbiotic plasmid (pRalta) of *C. taiwanensis*, was introduced into *R. solanacearum*,
509 resulting in a chimeric strain unable to nodulate *M. pudica,* the *C. taiwanensis* host. Subsequent
510 repeated inoculation trials on hundreds of *M. pudica* plants finally produced 3 rare nodules, from which
511 three distinct nodulating clones have been isolated (Marchetti et al., 2010). The nodulation efficiency of
512 these clones was however much lower than that of the natural *M. pudica* symbiont, *C. taiwanensis.*
513 These three first nodulating strains were then used as ancestors to evolve several parallel lineages of
514 bacteria through serial cycles of nodulation. At each cycle, nodules were collected, surface sterilized and
515 crushed to release nodule bacteria, which were directly re-inoculated onto new plants. These evolution
516 cycles roughly approximated what happens in nature, *i.e.* when nodules degenerate, bacteria are
517 released in the soil where they can survive as a free-living state or infect new plants. In a first
518 experiment, 16 cycles of nodulation on *M. pudica* were performed in 18 parallel lineages of evolved
519 populations along which the evolution of symbiotic properties was analyzed.
520

521 **3.2. Infection, a symbiotic trait drastically improved along lab-evolution**

522 Beside the acquisition of nodulation, a second crucial symbiotic property, intracellular infection of
523 nodules, was rapidly acquired. A first rudimentary level of intracellular infection was indeed observed in
524 nodules induced by two of the three nodulating ancestors as well as by evolved bacteria derived from
525 the third ancestor after few evolution cycles (Guan et al., 2013; Marchetti et al., 2010). These nodules
526 however always contained necrotic zones, indicative of plant defense reactions. Strikingly, this property
527 drastically improved within few additional cycles in a majority of parallel lineages. Nodules formed by
528 most final evolved clones were nicely invaded and devoid of necrosis (Marchetti et al., 2017; Marchetti
529 et al., 2014). Thus three distinct levels of infection were obtained during the course of this experiment,
530 extracellular, partially intracellular and nicely intracellular.

531

532 **3.3. Regulatory rewiring as a main adaptive change**

533 Resequencing the genomes of evolved bacteria and reconstructing the detected mutations allowed the
534 identification of adaptive mutations responsible for main symbiotic phenotypic shifts. The capacity to
535 nodulate was activated through the inactivation of the T3SS of *R. solanacearum* either via a stop
536 mutation in *hrcV*, an essential structural gene of the T3SS, or via stop mutations in the master regulator
537 of virulence *hrpG*, that not only controls the T3SS and effectors but also hundreds of other genes (Valls
538 et al., 2006). Rudimentary intracellular infection ability was gained through *hrpG* inactivation or via a
539 combination of the *hrcV* mutation and inactivation of various virulence regulators either *vsrA*, or *prhI* or
540 *hrpG* (Figure 2B) (Guan et al., 2013; Marchetti et al., 2010). Tremendous improvement of intracellular
541 infection together with drastic reduction of nodule necrosis then occurred via a rewiring of the virulence
542 regulatory network of *R. solanacearum* in several independent lineages. Mutations down-regulating
543 efpR, a virulence regulator and metabolic repressor, appeared five times independently to promote
544 infection. Repression of EfpR led to a constitutive induction of three intermediary genes encoding

545 membrane components (Rsc3146-3148) that in turn, likely indirectly, deregulate the expression of
546 hundreds of genes and allows the utilization of tens of new carbon, nitrogen and phosphorous
547 compounds (Figure 2BC) (Capela et al., 2017). In other lines, mutations that only slightly modulate the
548 responsiveness of *R. solanacearum* to cell density (*phcBQ* mutations) were sufficient to allow nice
549 intracellular infection (M. Tang, unpublished). Interestingly, both repression of EfpR and the delayed QS-
550 response modify the expression of hundreds of genes that many were in common and lead to a similar
551 metabolic profile (Figure 2C).
552 Thus, adaptation to symbiosis occurred via diverse genetic changes, either inactivation of structural
553 genes (*hrcV*), or inactivation of regulators (*hrpG*, *vsrA*, *prhI*, *efpR*) and recruitment of constitutively
554 expressed target genes (Rsc3146-3148), or modulation of regulatory paths to environmental signals
555 (*phcBQ*). The rewiring of the virulence regulatory network of *R. solanacearum* appeared as the main
556 adaptive mechanism to convert this pathogenic bacterium into legume symbionts, improving its
557 symbiotic infection ability while decreasing its pathogenicity. These results illustrated how modifications
558 in the regulatory circuitry of a bacterium can rapidly operate drastic changes in its lifestyle and
559 contribute to ecological transitions.
560
561 **3.4. Parallels between experimental and natural evolution of *Mimosa pudica* symbionts**
562 To evaluate to what extent such artificial evolution experiment can tell us about the natural evolutionary
563 history of legume symbionts, Clerissi *et al.* recently compared the emergence of *M. pudica* symbionts in
564 natural populations of *C. taiwanensis* with the experimental conversion of *R. solanacearum* into *Mimosa*
565 symbionts (Clerissi et al., 2018). In both cases the symbiotic transition was driven by the transfer of a
566 symbiotic plasmid. However the two processes displayed many fundamental differences in terms of
567 genetic backgrounds, *Ralstonia* and *Cupriavidus*, of time frame, more than 10 MY in one case and few
568 years in the case of lab-evolution, of environmental conditions, complex and changing in the first case

569 and simplified and constant in the second case, and finally of symbiosis achievement, mutualism was not
570 achieved in the evolution experiment while *C. taiwanensis* are effective and well adapted symbionts. Yet
571 a significant number of parallels were pointed out between the two evolutionary processes. First
572 symbiotic adaptation occurred on the recipient genomes and not on the symbiotic plasmid. Both
573 processes are also characterized by a dominance of purifying selection (*i.e.* a purge of non synonymous
574 mutations). This phenomenon may be due to an *imuABC* cassette encoding stress-responsive error-
575 prone DNA polymerases present in most extant strains of *C. taiwanensis* and generally present on the
576 symbiotic plasmid. In the evolution experiment, this cassette was shown to increase the mutation rate of
577 the recipient genome in the rhizosphere thus triggering a transient hypermutagenesis that accelerated
578 the adaptive evolution of the chimeric strain (Remigi et al., 2014). The co-transfer of this cassette
579 together with symbiosis genes was proposed to have enhanced the dissemination of symbiotic traits not
580 only in the *C. taiwanensis* branch but more broadly among α- and β-proteobacteria by facilitating the
581 post-HGT adaptations (Remigi et al., 2014). Finally, another interesting parallel was the co-option of the
582 Phc quorum-sensing system for symbiosis in both natural and experimental evolution. All together these
583 parallels demonstrated the relevance of experimental evolution approaches to potentially provide
584 important insights into natural evolutionary processes.
585

586 **Concluding remarks**

587 Rhizobia arose independently in distant bacterial lineages mostly following HGT of mutualistic traits.
588 Their parallel evolution from different ancestors accounts for the current diversity of the genetic
589 strategies these bacteria use to establish symbiosis with legumes. During evolution, extensive
590 recruitment, inactivation or modulation of indigenous functions together with deep regulatory rewiring
591 has achieved the ecological transition towards mutualism initiated by HGT. The success of this 'molecular
592 tinkering' essentially depends on the recipient genetic background as well as the chance a beneficial

593 mutation occurs. It also strongly relies on the strict selection process exerted by host plant that only
594 allows the entry of compatible variants into root tissues. The diversity of symbiotic molecular
595 mechanisms identified in rhizobia nicely illustrates the plasticity of bacterial genomes and the range of
596 possibilities to achieve a complex biological process. This also suggests that it would be possible to
597 evolve nitrogen-fixing symbiotic partners of non-legume crops engineered to nodulate.
598

599 **References**

600 Amadou, C., Pascal, G., Mangenot, S., Glew, M., Bontemps, C., Capela, D., et al. (2008). Genome
601 sequence of the beta-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome*
602 *Res,* 18, 1472-1483.
603 Andrews, M., De Meyer, S., James, E.K., Stępkowski, T., Hodge, S., Simon, M.F., et al. (2018). Horizontal
604 transfer of symbiosis genes within and between rhizobial genera: occurrence and importance. *Genes*
605 *(Basel),* 9, 321.
606 Aoki, S., Ito, M., and Iwasaki, W. (2013). From beta- to alpha-proteobacteria: the origin and evolution of
607 rhizobial nodulation genes *nodIJ*. *Mol Biol Evol,* 30, 2494-2508.
608 Arnold, M.F.F., Penterman, J., Shabab, M., Chen, E.J., and Walker, G.C. (2018). Important late-stage
609 symbiotic role of the *Sinorhizobium meliloti* exopolysaccharide succinoglycan. *J Bacteriol*, 200, e00665-
610 17.
611 Barcellos, F.G., Menna, P., Batista, J.S.D., and Hungria, M. (2007). Evidence of horizontal transfer of
612 symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs
613 *Sinorhizobium* (*Ensifer*) *fredii* and *Bradyrhizobium elkanii* in a Brazilian Savannah soil. *Appl Environ*
614 *Microbiol,* 73, 2635-2643.
615 Barnett, M.J., and Long, S.R. (2018). Novel genes and regulators that influence production of cell surface
616 exopolysaccharides in *Sinorhizobium meliloti*. *J Bacteriol*, 200, e00501-17.

617 Bartsev, A.V., Deakin, W.J., Boukli, N.M., McAlvin, C.B., Stacey, G., Malnoe, P., et al. (2004). NopL, an
618 effector protein of *Rhizobium* sp. NGR234, thwarts activation of plant defense reactions. *Plant Physiol*,
619 134, 871-879.
620 Battisti, L., Lara, J.C., and Leigh, J.A. (1992). Specific oligosaccharide form of the *Rhizobium meliloti*
621 exopolysaccharide promotes nodule invasion in alfalfa. *Proc Natl Acad Sci U S A,* 89, 5625-5629.
622 Bauer, E., Kaspar, T., Fischer, H.M., and Hennecke, H. (1998). Expression of the *fixR-nifA* operon in
623 *Bradyrhizobium japonicum* depends on a new response regulator, RegR. *J Bacteriol*, 180, 3853-3863.
624 Bañuelos-Vazquez, L.A., Torres Tejerizo, G., and Brom, S. (2017). Regulation of conjugative transfer of
625 plasmids and integrative conjugative elements. *Plasmid*, 91, 82-89.
626 Bhagwat, A.A., Gross, K.C., Tully, R.E., and Keister, D.L. (1996). Beta-glucan synthesis in *Bradyrhizobium*
627 *japonicum*: characterization of a new locus (*ndvC*) influencing beta-(1-->6) linkages. *J Bacteriol*, 178,
628 4635-4642.
629 Bhagwat, A.A., Tully, R.E., and Keister, D.L. (1992). Isolation and characterization of an *ndvB* locus from
630 *Rhizobium fredii*. *Mol Microbiol,* 6, 2159-2165.
631 Black, M., Moolhuijzen, P., Chapman, B., Barrero, R., Howieson, J., Hungria, M., et al. (2012). The genetics
632 of symbiotic nitrogen fixation: comparative genomics of 14 rhizobia strains by resolution of protein
633 clusters. *Genes (Basel)*, 3, 138-166.
634 Bladergroen, M.R., Badelt, K., and Spaink, H.P. (2003). Infection-blocking genes of a symbiotic *Rhizobium*
635 *leguminosarum* strain that are involved in temperature-dependent protein secretion. *Mol Plant-Microbe*
636 *Interact*, 16, 53-64.
637 Bobik, C., Meilhoc, E., and Batut, J. (2006). FixJ: a major regulator of the oxygen limitation response and
638 late symbiotic functions of *Sinorhizobium meliloti*. *J Bacteriol*, 188, 4890-4902.
639 Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., et al. (2011). Nodulation of
640 *Aeschynomene afraspera* and *A. indica* by photosynthetic *Bradyrhizobium* sp. strain ORS285: The Nod-

641 dependent Versus the Nod-independent symbiotic interaction. *Mol Plant-Microbe Interact*, 24, 1359-
642 1371.
643 Bonomi, H.R., Posadas, D.M., Paris, G., Carrica, M.e.C., Frederickson, M., Pietrasanta, L.I., et al. (2012).
644 Light regulates attachment, exopolysaccharide production, and nodulation in *Rhizobium leguminosarum*
645 through a LOV-histidine kinase photoreceptor. *Proc Natl Acad Sci U S A*, 109, 12135-12140.
646 Breedveld, M.W., and Miller, K.J. (1994). Cyclic beta-glucans of members of the family Rhizobiaceae.
647 *Microbiol Rev*, 58, 145-161.
648 Broughton, W.J., Hanin, M., Relic, B., Kopciñska, J., Golinowski, W., Simsek, S., et al. (2006). Flavonoid-
649 inducible modifications to rhamnan O antigens are necessary for *Rhizobium* sp. strain NGR234-legume
650 symbioses. *J Bacteriol*, 188, 3654-3663.
651 Bélanger, L., Dimmick, K.A., Fleming, J.S., and Charles, T.C. (2009). Null mutations in *Sinorhizobium*
652 *meliloti exoS* and *chvI* demonstrate the importance of this two-component regulatory system for
653 symbiosis. *Mol Microbiol*, 74, 1223-1237.
654 Calatrava-Morales, N., McIntosh, M., and Soto, M.J. (2018). Regulation mediated by N-acyl homoserine
655 lactone quorum sensing signals in the *Rhizobium*-Legume symbiosis. *Genes (Basel)*, 9, 263.
656 Calderón-Flores, A., Du Pont, G., Huerta-Saquero, A., Merchant-Larios, H., Servín-González, L., and Durán,
657 S. (2005). The stringent response is required for amino acid and nitrate utilization, nod factor regulation,
658 nodulation, and nitrogen fixation in *Rhizobium etli*. *J Bacteriol*, 187, 5075-5083.
659 Campbell, G.R.O., Reuhs, B.L., and Walker, G.C. (2002). Chronic intracellular infection of alfalfa nodules
660 by *Sinorhizobium meliloti* requires correct lipopolysaccharide core. *Proc Natl Acad Sci U S A*, 99, 3938-
661 3943.
662 Capela, D., Carrere, S., and Batut, J. (2005). Transcriptome-based identification of the *Sinorhizobium*
663 *meliloti* NodD1 regulon. *Appl Environ Microbiol,* 71, 4910-4913.

664 Capela, D., Marchetti, M., Clerissi, C., Perrier, A., Guetta, D., Gris, C., et al. (2017). Recruitment of a
665 lineage-specific virulence regulatory pathway promotes intracellular infection by a plant pathogen
666 experimentally evolved into a legume symbiont. *Mol Biol Evol,* 34, 2503-2521.
667 Capoen, W., Goormachtig, S., and Holsters, M. (2010). Water-tolerant legume nodulation. *J Exp Bot,* 61,
668 1251-1255.
669 Cerri, M.R., Frances, L., Kelner, A., Fournier, J., Middleton, P.H., Auriac, M.C., et al. (2016). The symbiosis-
670 related ERN transcription factors act in concert to coordinate rhizobial host root infection. *Plant Physiol,*
671 171, 1037-1054.
672 Cerri, M.R., Wang, Q., Stolz, P., Folgmann, J., Frances, L., Katzer, K., et al. (2017). The ERN1 transcription
673 factor gene is a target of the CCaMK/CYCLOPS complex and controls rhizobial infection in *Lotus*
674 *japonicus*. *New Phytol*, 215, 323-337.
675 Chaintreuil, C., Perrier, X., Martin, G., Fardoux, J., Lewis, G.P., Brottier, L., et al. (2018). Naturally
676 occurring variations in the *nod*-independent model legume *Aeschynomene evenia* and relatives: a
677 resource for nodulation genetics. *BMC Plant Biol*, 18, 54.
678 Chen, W.M., Moulin, L., Bontemps, C., Vandamme, P., Bena, G., and Boivin-Masson, C. (2003). Legume
679 symbiotic nitrogen fixation by beta-proteobacteria is widespread in nature. *J Bacteriol*, 185, 7266-7272.
680 Cheng, H.P., and Walker, G.C. (1998). Succinoglycan is required for initiation and elongation of infection
681 threads during nodulation of alfalfa by *Rhizobium meliloti*. *J Bacteriol*, 180, 5183-5191.
682 Clerissi, C., Touchon, M., Capela, D., Tang, M., Cruveiller, S., Parker, M.A., et al. (2018). Parallels between
683 experimental and natural evolution of legume symbionts. *Nat Commun*, 9, 2264.
684 Cosseau, C., and Batut, J. (2004). Genomics of the *ccoNOQP*-encoded *cbb*(3) oxidase complex in bacteria.
685 *Arch Microbiol*, 181, 89-96.

686 Coutinho, B.G., Mitter, B., Talbi, C., Sessitsch, A., Bedmar, E.J., Halliday, N., et al. (2013). Regulon studies
687 and in planta role of the BraI/R quorum-sensing system in the plant-beneficial *Burkholderia* cluster. *Appl*
688 *Environ Microbiol,* 79, 4421-4432.
689 Crespo-Rivas, J.C., Margaret, I., Hidalgo, A., Buendía-Clavería, A.M., Ollero, F.J., López-Baena, F.J., et al.
690 (2009). *Sinorhizobium fredii* HH103 *cgs* mutants are unable to nodulate determinate- and indeterminate
691 nodule-forming legumes and overproduce an altered EPS. *Mol Plant Microbe Interact*, 22, 575-588.
692 Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., et al. (2015). Convergent evolution
693 of endosymbiont differentiation in Dalbergioid and inverted repeat-lacking clade legumes mediated by
694 nodule-specific cysteine-rich peptides. *Plant Physiol*, 169, 1254-1265.
695 D'Antuono, A.L., Ott, T., Krusell, L., Voroshilova, V., Ugalde, R.A., Udvardi, M., et al. (2008). Defects in
696 rhizobial cyclic glucan and lipopolysaccharide synthesis alter legume gene expression during nodule
697 development. *Mol Plant Microbe Interact*, 21, 50-60.
698 D'Haeze, W., and Holsters, M. (2002). Nod factor structures, responses, and perception during initiation
699 of nodule development. *Glycobiology*, 12, 79R-105R.
700 D'Haeze, W., and Holsters, M. (2004). Surface polysaccharides enable bacteria to evade plant immunity.
701 *Trends Microbiol*, 12, 555-561.
702 Daniels, R., De Vos, D.E., Desair, J., Raedschelders, G., Luyten, E., Rosemeyer, V., et al. (2002). The *cin*
703 quorum sensing locus of *Rhizobium etli* CNPAF512 affects growth and symbiotic nitrogen fixation. *J Biol*
704 *Chem*, 277, 462-468.
705 Danino, V.E., Wilkinson, A., Edwards, A., and Downie, J.A. (2003). Recipient-induced transfer of the
706 symbiotic plasmid pRL1JI in *Rhizobium leguminosarum* bv. *viciae* is regulated by a quorum-sensing relay.
707 *Mol Microbiol*, 50, 511-525.
708 Davies, B.W., and Walker, G.C. (2007). Identification of novel *Sinorhizobium meliloti* mutants
709 compromised for oxidative stress protection and symbiosis. *J Bacteriol*, 189, 2110-2113.

710 De Meyer, S.E., Briscoe, L., Martínez-Hidalgo, P., Agapakis, C.M., de-Los Santos, P.E., Seshadri, R., et al.
711 (2016). Symbiotic *Burkholderia* species show diverse arrangements of *nif/fix* and *nod* genes and lack
712 typical high-affinity cytochrome *cbb*3 oxidase genes. *Mol Plant Microbe Interact*, 29, 609-619.
713 Deakin, W.J., and Broughton, W.J. (2009). Symbiotic use of pathogenic strategies: rhizobial protein
714 secretion systems. *Nat Rev Microbiol*, 7, 312-320.
715 del Cerro, P., Rolla-Santos, A.A., Gomes, D.F., Marks, B.B., del Rosario Espuny, M., Rodríguez-Carvajal, M.,
716 et al. (2015). Opening the "black box" of *nodD3*, *nodD4* and *nodD5* genes of *Rhizobium tropici* strain CIAT
717 899. *BMC Genomics*, 16, 864.
718 Delavat, F., Miyazaki, R., Carraro, N., Pradervand, N., and van der Meer, J.R. (2017). The hidden life of
719 integrative and conjugative elements. *FEMS Microbiol Rev*, 41, 512-537.
720 Ding, H., and Hynes, M.F. (2009). Plasmid transfer systems in the rhizobia. *Can J Microbiol,* 55, 917-927.
721 Dixon, R., and Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nat Rev Microbiol*, 2,
722 621-631.
723 Downie, J.A. (2010). The roles of extracellular proteins, polysaccharides and signals in the interactions of
724 rhizobia with legume roots. *FEMS Microbiol Rev,* 34, 150-170.
725 Dusha, I., and Kondorosi, A. (1993). Genes at different regulatory levels are required for the ammonia
726 control of nodulation in *Rhizobium meliloti*. *Mol Gen Genet*, 240, 435-444.
727 Dylan, T., Nagpal, P., Helinski, D.R., and Ditta, G.S. (1990). Symbiotic pseudorevertants of *Rhizobium*
728 *meliloti ndv* mutants. *J Bacteriol,* 172, 1409-1417.
729 Estrada-de Los Santos, P., Palmer, M., Chávez-Ramírez, B., Beukes, C., Steenkamp, E.T., Briscoe, L., et al.
730 (2018). Whole genome analyses suggests that *Burkholderia sensu lato* contains two additional novel
731 genera (*Mycetohabitans* gen. nov., and *Trinickia* gen. nov.): implications for the evolution of diazotrophy
732 and nodulation in the Burkholderiaceae. *Genes (Basel),* 9, 389.

733 Fabre, S., Gully, D., Poitout, A., Patrel, D., Arrighi, J.F., Giraud, E., et al. (2015). Nod Factor-independent
734 nodulation in *Aeschynomene evenia* required the common plant-microbe symbiotic toolkit. *Plant Physiol,*
735 169, 2654-2664.
736 Fan, Y., Liu, J., Lyu, S., Wang, Q., Yang, S., and Zhu, H. (2017). The soybean *rfg1* gene restricts nodulation
737 by *Sinorhizobium fredii* USDA193. *Front Plant Sci*, 8, 1548.
738 Fellay, R., Hanin, M., Montorzi, G., Frey, J., Freiberg, C., Golinowski, W., et al. (1998). *nodD2* of *Rhizobium*
739 sp. NGR234 is involved in the repression of the *nodABC* operon. *Mol Microbiol*, 27, 1039-1050.
740 Fraysse, N., Couderc, F., and Poinsot, V. (2003). Surface polysaccharide involvement in establishing the
741 rhizobium-legume symbiosis. *Eur J Biochem*, 270, 1365-1380.
742 Gage, D.J. (2002). Analysis of infection thread development using Gfp- and DsRed-expressing
743 *Sinorhizobium meliloti*. *J Bacteriol*, 184, 7042-7046.
744 Galibert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., et al. (2001). The composite genome
745 of the legume symbiont *Sinorhizobium meliloti*. *Science*, 293, 668-672.
746 Garcia, M., Dunlap, J., Loh, J., and Stacey, G. (1996). Phenotypic characterization and regulation of the
747 *nolA* gene of *Bradyrhizobium japonicum*. *Mol Plant Microbe Interact*, 9, 625-636.
748 Geddes, B.A., González, J.E., and Oresnik, I.J. (2014). Exopolysaccharide production in response to
749 medium acidification is correlated with an increase in competition for nodule occupancy. *Mol Plant*
750 *Microbe Interact*, 27, 1307-1317.
751 Gherbi, H., Markmann, K., Svistoonoff, S., Estevan, J., Autran, D., Giczey, G., et al. (2008). SymRK defines
752 a common genetic basis for plant root endosymbioses with arbuscular mycorrhiza fungi, rhizobia, and
753 Frankia bacteria. *Proc Natl Acad Sci U S A*, 105, 4928-4932.
754 Giraud, E., Hannibal, L., Fardoux, L., Vermeglio, A., and Dreyfus, B. (2000). Effect of *Bradyrhizobium*
755 photosynthesis on stem nodulation of *Aeschynomene sensitiva*. *Proc Natl Acad Sci U S A*, 97, 14795-
756 14800.

757 Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.C., et al. (2007). Legumes symbioses:
758 absence of *nod* genes in photosynthetic bradyrhizobia. *Science*, 316, 1307-1312.
759 Gough, C., and Cullimore, J. (2011). Lipo-chitooligosaccharide signaling in endosymbiotic plant-microbe
760 interactions. *Mol Plant-Microbe Interact*, 24, 867-878.
761 Guan, S.H., Gris, C., Cruveiller, S., Pouzet, C., Tasse, L., Leru, A., et al. (2013). Experimental evolution of
762 nodule intracellular infection in legume symbionts. *ISME J*, 7, 1367-1377.
763 Guefrachi, I., Pierre, O., Timchenko, T., Alunni, B., Barrière, Q., Czernic, P., et al. (2015). *Bradyrhizobium*
764 BclA Is a peptide transporter required for bacterial differentiation in symbiosis with *Aeschynomene*
765 legumes. *Mol Plant Microbe Interact*, 28, 1155-1166.
766 Gully, D., Czernic, P., Cruveiller, S., Mahé, F., Longin, C., Vallenet, D., et al. (2018). Transcriptome profiles
767 of Nod Factor-independent symbiosis in the tropical legume *Aeschynomene evenia*. *Sci Rep*, 8, 10934.
768 Gurich, N., and González, J.E. (2009). Role of quorum sensing in *Sinorhizobium meliloti*-Alfalfa symbiosis.
769 *J Bacteriol*, 191, 4372-4382.
770 Gyaneshwar, P., Hirsch, A.M., Moulin, L., Chen, W.M., Elliott, G.N., Bontemps, C., et al. (2011). Legume-
771 nodulating Betaproteobacteria: diversity, host range, and future prospects. *Mol Plant-Microbe Interact*,
772 24, 1276-1288.
773 Haag, A.F., Baloban, M., Sani, M., Kerscher, B., Pierre, O., Farkas, et al. (2011). Protection of
774 *Sinorhizobium* against host cysteine-rich antimicrobial peptides is critical for symbiosis. *PLoS Biol*, 9,
775 e1001169.
776 Hakoyama, T., Niimi, K., Watanabe, H., Tabata, R., Matsubara, J., Sato, S., et al. (2009). Host plant
777 genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. *Nature*, 462, 514-517.
778 Hauser, F., Pessi, G., Friberg, M., Weber, C., Rusca, N., Lindemann, A., et al. (2007). Dissection of the
779 *Bradyrhizobium japonicum* NifA+sigma54 regulon, and identification of a ferredoxin gene (*fdxN*) for
780 symbiotic nitrogen fixation. *Mol Genet Genomics*, 278, 255-271.

781 Heavner, M.E., Qiu, W.-G., and Cheng, H.-P. (2015). Phylogenetic co-occurrence of ExoR, ExoS, and ChvI,
782 components of the RSI bacterial invasion switch, suggests a key adaptive mechanism regulating the
783 transition between free-living and host-invading phases in Rhizobiales. *PLoS One*, 10, e0135655.
784 Hirsch, A.M., Wilson, K.J., Jones, J.D.G., Bang, M., Walker, V.V., and Ausubel, F.M. (1984). *Rhizobium*
785 *meliloti* nodulation genes allow *Agrobacterium tumefaciens* and *Escherichia coli* to form pseudonodules
786 on alfalfa. *J Bacteriol*, 158, 1133-1143.
787 Hoang, H.H., Becker, A., and González, J.E. (2004). The LuxR homolog ExpR, in combination with the Sin
788 quorum sensing system, plays a central role in *Sinorhizobium meliloti* gene expression. *J Bacteriol*, 186,
789 5460-5472.
790 Hocher, V., Alloisio, N., Bogusz, D., and Normand, P. (2011). Early signaling in actinorhizal symbioses.
791 *Plant Signal Behav*, 6, 1377-1379.
792 Hubber, A., Vergunst, A.C., Sullivan, J.T., Hooykaas, P.J.J., and Ronson, C.W. (2004). Symbiotic
793 phenotypes and translocated effector proteins of the *Mesorhizobium loti* strain R7A VirB/D4 type IV
794 secretion system. *Mol Microbiol*, 54, 561-574.
795 Hubber, A.M., Sullivan, J.T., and Ronson, C.W. (2007). Symbiosis-induced cascade regulation of the
796 *Mesorhizobium loti* R7A VirB/D4 type IV secretion system. *Mol Plant-Microbe Interact*, 20, 255-261.
797 Ingram-Smith, C., and Miller, K.J. (1998). Effects of ionic and osmotic strength on the glucosyltransferase
798 of *Rhizobium meliloti* responsible for cyclic beta-(1,2)-glucan biosynthesis. *Appl Environ Microbiol*, 64,
799 1290-1297.
800 Irving, S.E., and Corrigan, R.M. (2018). Triggering the stringent response: signals responsible for
801 activating (p)ppGpp synthesis in bacteria. *Microbiology*, 164, 268-276.
802 Janczarek, M. (2011). Environmental signals and regulatory pathways that influence exopolysaccharide
803 production in rhizobia. *Int J Mol Sci*, 12, 7898-7933.

804 Jaszek, M., Janczarek, M., Piersiak, T., and Grzywnowicz, K. (2014). The response of the *Rhizobium*
805 *leguminosarum* bv. *trifolii* wild-type and exopolysaccharide-deficient mutants to oxidative stress. *Plant*
806 *Soil*, 376, 75-94.
807 Jones, K.M., Sharopova, N., Lohar, D.P., Zhang, J.Q., VandenBosch, K.A., and Walker, G.C. (2008).
808 Differential response of the plant *Medicago truncatula* to its symbiont *Sinorhizobium meliloti* or an
809 exopolysaccharide-deficient mutant. *Proc Natl Acad Sci U S A*, 105, 704-709.
810 Jourand, P., Renier, A., Rapior, S., de Faria, S.M., Prin, Y., Galiana, A., et al. (2005). Role of methylotrophy
811 during symbiosis between *Methylobacterium nodulans* and *Crotalaria podocarpa*. *Mol Plant-Microbe*
812 *Interact*, 18, 1061-1068.
813 Kaminski, P.A., and Elmerich, C. (1998). The control of *Azorhizobium caulinodans nifA* expression by
814 oxygen, ammonia and by the HF-I-like protein, NrfA. *Mol Microbiol*, 28, 603-613.
815 Karmakar, K., Kundu, A., Rizvi, A.Z., Dubois, E., Severac, D., Czernic, P., et al. (2019). Transcriptomic
816 analysis with the progress of symbiosis in 'crack-entry' legume *Arachis hypogaea* highlights its contrast
817 with 'infection thread' adapted legumes. *Mol Plant Microbe Interact*, MPMI06180174R.
818 Kawaharada, Y., Kelly, S., Nielsen, M.W., Hjuler, C.T., Gysel, K., Muszyński, A., et al. (2015). Receptor-
819 mediated exopolysaccharide perception controls bacterial infection. *Nature*, 523, 308-312.
820 Kawaharada, Y., Nielsen, M.W., Kelly, S., James, E.K., Andersen, K.R., Rasmussen, S.R., et al. (2017).
821 Differential regulation of the Epr3 receptor coordinates membrane-restricted rhizobial colonization of
822 root nodule primordia. *Nat Commun*, 8, 14534.
823 Kelly, S., Sullivan, J.T., Kawaharada, Y., Radutoiu, S., Ronson, C.W., and Stougaard, J. (2018). Regulation of
824 Nod factor biosynthesis by alternative NodD proteins at distinct stages of symbiosis provides additional
825 compatibility scrutiny. *Environ Microbiol*, 20, 97-110.

826 Kelly, S.J., Muszyński, A., Kawaharada, Y., Hubber, A.M., Sullivan, J.T., Sandal, N., et al. (2013).
827 Conditional requirement for exopolysaccharide in the *Mesorhizobium-Lotus* symbiosis. *Mol Plant*
828 *Microbe Interact*, 26, 319-329.
829 Kim, C.H., Tully, R.E., and Keister, D.L. (1989). Exopolysaccharide-deficient mutants of *Rhizobium fredii*
830 HH303 which are symbiotically effective. *Appl Environ Microbiol*, 55, 1852-1854.
831 Kiss, E., Kereszt, A., Barta, F., Stephens, S., Reuhs, B.L., Kondorosi, A., et al. (2001). The *rkp*-3 gene region
832 of *Sinorhizobium meliloti* Rm41 contains strain-specific genes that determine K antigen structure. *Mol*
833 *Plant Microbe Interact*, 14, 1395-1403.
834 Kobayashi, H., Graven, Y.N., Broughton, W.J., and Perret, X. (2004). Flavonoids induce temporal shifts in
835 gene-expression of *nod*-box controlled loci in *Rhizobium* sp. NGR234. *Mol Microbiol*, 51, 335-347.
836 Kouchi, H., Imaizumi-Anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., et al. (2010).
837 How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. *Plant Cell*
838 *Physiol*, 51, 1381-1397.
839 Krause, A., Doerfel, A., and Göttfert, M. (2002). Mutational and transcriptional analysis of the type III
840 secretion system of *Bradyrhizobium japonicum*. *Mol Plant Microbe Interact*, 15, 1228-1235.
841 Krysciak, D., Grote, J., Rodriguez Orbegoso, M., Utpatel, C., Förstner, K.U., Li, L., et al. (2014). RNA
842 sequencing analysis of the broad-host-range strain *Sinorhizobium fredii* NGR234 identifies a large set of
843 genes linked to quorum sensing-dependent regulation in the background of a *traI* and *ngrI* deletion
844 mutant. *Appl Environ Microbiol*, 80, 5655-5671.
845 Ktari, A., Nouioui, I., Furnholm, T., Swanson, E., Ghodhbane-Gtari, F., Tisa, L.S., et al. (2017). Permanent
846 draft genome sequence of *Frankia* sp. NRRL B-16219 reveals the presence of canonical *nod* genes, which
847 are highly homologous to those detected in Candidatus *Frankia* Dg1 genome. *Stand Genomic Sci*, 12, 51.

848 Laguerre, G., Nour, S.M., Macheret, V., Sanjuan, J., Drouin, P., and Amarger, N. (2001). Classification of
849 rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus*
850 *vulgaris* symbionts. *Microbiology*, 147, 981-993.
851 Ledbetter, R.N., Garcia Costas, A.M., Lubner, C.E., Mulder, D.W., Tokmina-Lukaszewska, M., Artz, J.H., et
852 al. (2017). The electron bifurcating FixABCX protein complex from *Azotobacter vinelandii*: generation of
853 low-potential reducing equivalents for nitrogenase catalysis. *Biochemistry*, 56, 4177-4190.
854 Lehman, A.P., and Long, S.R. (2013). Exopolysaccharides from *Sinorhizobium meliloti* can protect against
855 H_2O_2-dependent damage. *J Bacteriol*, 195, 5362-5369.
856 Lercher, M.J., and Pál, C. (2008). Integration of horizontally transferred genes into regulatory interaction
857 networks takes many million years. *Mol Biol Evol*, 25, 559-567.
858 Ling, J., Wang, H., Wu, P., Li, T., Tang, Y., Naseer, N., et al. (2016). Plant nodulation inducers enhance
859 horizontal gene transfer of *Azorhizobium caulinodans* symbiosis island. *Proc Natl Acad Sci U S A*, 113,
860 13875-13880.
861 Liu, W., Sun, Y., Shen, R., Dang, X., Liu, X., Sui, F., et al. (2018). A chemotaxis-like pathway of
862 *Azorhizobium caulinodans* controls flagella-driven motility, which regulates biofilm formation,
863 exopolysaccharide biosynthesis, and competitive nodulation. *Mol Plant Microbe Interact*, 31, 737-749.
864 Loh, J., and Stacey, G. (2003). Nodulation gene regulation in *Bradyrhizobium japonicum*: a unique
865 integration of global regulatory circuits. *Appl Environ Microbiol*, 69, 10-17.
866 Loh, J.T., Yuen-Tsai, J.P., Stacey, M.G., Lohar, D., Welborn, A., and Stacey, G. (2001). Population density-
867 dependent regulation of the *Bradyrhizobium japonicum* nodulation genes. *Mol Microbiol*, 42, 37-46.
868 López-Baena, F.J., Ruiz-Sainz, J.E., Rodríguez-Carvajal, M.A., and Vinardell, J.M. (2016). Bacterial
869 molecular signals in the *Sinorhizobium fredii*-Soybean symbiosis. *Int J Mol Sci*, 17.
870 MacLean, A.M., Finan, T.M., and Sadowsky, M.J. (2007). Genomes of the symbiotic nitrogen-fixing
871 bacteria of legumes. *Plant Physiology*, 144, 615-622.

872 Madsen, L.H., Tirichine, L., Jurkiewicz, A., Sullivan, J.T., Heckmann, A.B., Bek, A.S., et al. (2010). The
873 molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*.
874 *Nat Commun*, 1, 10.
875 Marchetti, M., Capela, D., Glew, M., Cruveiller, S., Chane-Woon-Ming, B., Gris, C., et al. (2010).
876 Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biol*, 8, e1000280.
877 Marchetti, M., Clerissi, C., Yousfi, Y., Gris, C., Bouchez, O., Rocha, E., et al. (2017). Experimental evolution
878 of rhizobia may lead to either extra- or intracellular symbiotic adaptation depending on the selection
879 regime. *Mol Ecol*, 26, 1818-1831.
880 Marchetti, M., Jauneau, A., Capela, D., Remigi, P., Gris, C., Batut, J., et al. (2014). Shaping bacterial
881 symbiosis with legumes by experimental evolution. *Mol Plant-Microbe Interact*, 27, 956-964.
882 Marie, C., Deakin, W.J., Viprey, V., Kopciñska, J., Golinowski, W., Krishnan, H.B., et al. (2003).
883 Characterization of Nops, nodulation outer proteins, secreted via the type III secretion system of
884 NGR234. *Mol Plant Microbe Interact*, 16, 743-751.
885 Maruya, J., and Saeki, K. (2010). The bacA gene homolog, mlr7400, in *Mesorhizobium loti* MAFF303099 is
886 dispensable for symbiosis with *Lotus japonicus* but partially capable of supporting the symbiotic function
887 of *bacA* in *Sinorhizobium meliloti*. *Plant Cell Physiol*, 51, 1443-1452.
888 Masson-Boivin, C., Giraud, E., Perret, X., and Batut, J. (2009). Establishing nitrogen-fixing symbiosis with
889 legumes: how many rhizobium recipes? *Trends Microbiol*, 17, 458-466.
890 Masson-Boivin, C., and Sachs, J.L. (2018). Symbiotic nitrogen fixation by rhizobia-the roots of a success
891 story. *Curr Opin Plant Biol*, 44, 7-15.
892 Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., et al. (2003). A novel
893 family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small,
894 secreted polypeptides with conserved cysteine motifs. *Plant Physiol*, 132, 161-173.

895 Miller-Williams, M., Loewen, P.C., and Oresnik, I.J. (2006). Isolation of salt-sensitive mutants of
896 *Sinorhizobium meliloti* strain Rm1021. *Microbiology*, 152, 2049-2059.
897 Moris, M., Braeken, K., Schoeters, E., Verreth, C., Beullens, S., Vanderleyden, J., et al. (2005). Effective
898 symbiosis between *Rhizobium etli* and *Phaseolus vulgaris* requires the alarmone ppGpp. *J Bacteriol*, 187,
899 5460-5469.
900 Morris, J., and González, J.E. (2009). The novel genes *emmABC* are associated with exopolysaccharide
901 production, motility, stress adaptation, and symbiosis in *Sinorhizobium meliloti*. *J Bacteriol*, 191, 5890-
902 5900.
903 Mousavi, S.A., Willems, A., Nesme, X., de Lajudie, P., and Lindström, K. (2015). Revised phylogeny of
904 Rhizobiaceae: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations.
905 *Syst Appl Microbiol*, 38, 84-90.
906 Muszyński, A., Heiss, C., Hjuler, C.T., Sullivan, J.T., Kelly, S.J., Thygesen, M.B., et al. (2016). Structures of
907 exopolysaccharides involved in receptor-mediated perception of *Mesorhizobium loti* by *Lotus japonicus*. *J*
908 *Biol Chem*, 291, 20946-20961.
909 Mølbak, L., Molin, S., and Kroer, N. (2007). Root growth and exudate production define the frequency of
910 horizontal plasmid transfer in the rhizosphere. *FEMS Microbiol Ecol*, 59, 167-176.
911 Nandasena, K.G., O'Hara, G.W., Tiwari, R.P., Sezmis, E., and Howieson, J.G. (2007). In situ lateral transfer
912 of symbiosis islands results in rapid evolution of diverse competitive strains of mesorhizobia suboptimal
913 in symbiotic nitrogen fixation on the pasture legume *Biserrula pelecinus* L. *Environ Microbiol*, 9, 2496-
914 2511.
915 Nelson, M.S., Chun, C.L., and Sadowsky, M.J. (2017). Type IV effector proteins involved in the *Medicago-*
916 *Sinorhizobium* Symbiosis. *Mol Plant Microbe Interact*, 30, 28-34.
917 Nelson, M.S., and Sadowsky, M.J. (2015). Secretion systems and signal exchange between nitrogen-fixing
918 rhizobia and legumes. *Front Plant Sci*, 6, 491.

919 Nishihata, S., Kondo, T., Tanaka, K., Ishikawa, S., Takenaka, S., Kang, C.M., et al. (2018). *Bradyrhizobium*
920 *diazoefficiens* USDA110 PhaR functions for pleiotropic regulation of cellular processes besides PHB
921 accumulation. *BMC Microbiol*, 18, 156.
922 Noel, K.D., Box, J.M., and Bonne, V.J. (2004). 2-O-methylation of fucosyl residues of a rhizobial
923 lipopolysaccharide is increased in response to host exudate and is eliminated in a symbiotically defective
924 mutant. *Appl Environ Microbiol*, 70, 1537-1544.
925 Noel, K.D., Forsberg, L.S., and Carlson, R.W. (2000). Varying the abundance of O antigen in *Rhizobium etli*
926 and its effect on symbiosis with *Phaseolus vulgaris*. *J Bacteriol*, 182, 5317-5324.
927 Nogales, J., Blanca-Ordóñez, H., Olivares, J., and Sanjuán, J. (2013). Conjugal transfer of the
928 *Sinorhizobium meliloti* 1021 symbiotic plasmid is governed through the concerted action of one- and
929 two-component signal transduction regulators. *Environ Microbiol*, 15, 811-821.
930 Nouwen, N., Arrighi, J.F., Cartieaux, F., Chaintreuil, C., Gully, D., Klopp, C., et al. (2017). The role of
931 rhizobial (NifV) and plant (FEN1) homocitrate synthases in *Aeschynomene*/photosynthetic
932 *Bradyrhizobium* symbiosis. *Sci Rep*, 7, 448.
933 Ojeda, K.J., Box, J.M., and Noel, K.D. (2010). Genetic basis for Rhizobium etli CE3 O-antigen O-methylated
934 residues that vary according to growth conditions. *J Bacteriol*, 192, 679-690.
935 Okazaki, S., Kaneko, T., Sato, S., and Saeki, K. (2013). Hijacking of leguminous nodulation signaling by the
936 rhizobial type III secretion system. *Proc Natl Acad Sci U S A*, 110, 17131-17136.
937 Okazaki, S., Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., et al. (2016). Rhizobium-
938 legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. *ISME J*,
939 10, 64-74.
940 Okazaki, S., Zehner, S., Hempel, J., Lang, K., and Göttfert, M. (2009). Genetic organization and functional
941 analysis of the type III secretion system of *Bradyrhizobium elkanii*. *FEMS Microbiol Lett*, 295, 88-95.

942 Okubo, T., Fukushima, S., and Minamisawa, K. (2012). Evolution of *Bradyrhizobium-Aeschynomene*
943 mutualism: living testimony of the ancient world or highly evolved state? *Plant Cell Physiol*, 53, 2000-
944 2007.
945 Oldroyd, G.E. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic
946 associations in plants. *Nat Rev Microbiol*, 11, 252-263.
947 Parada, M., Vinardell, J.M., Ollero, F.J., Hidalgo, A., Gutiérrez, R., Buendía-Clavería, A.M., et al. (2006).
948 *Sinorhizobium fredii* HH103 mutants affected in capsular polysaccharide (KPS) are impaired for
949 nodulation with soybean and *Cajanus cajan*. *Mol Plant Microbe Interact*, 19, 43-52.
950 Ormeño-Orrillo, E., Servín-Garcidueñas, L.E., Rogel, M.A., González, V., Peralta, H., Mora, J., et al. (2015).
951 Taxonomy of rhizobia and agrobacteria from the Rhizobiaceae family in light of genomics. *Syst Appl*
952 *Microbiol*, 38, 287-291.
953 Pellock, B.J., Cheng, H.P., and Walker, G.C. (2000). Alfalfa root nodule invasion efficiency is dependent on
954 *Sinorhizobium meliloti* polysaccharides. *J Bacteriol*, 182, 4310-4318.
955 Perret, X., Staehelin, C., and Broughton, W.J. (2000). Molecular basis of symbiotic promiscuity. *Microbiol*
956 *Mol Biol Rev*, 64, 180-201.
957 Persson, T., Battenberg, K., Demina, I.V., Vigil-Stenman T., Vanden Heuvel B., Pujic P., et al. (2015).
958 Candidatus *Frankia Datiscae* Dg1, the actinobacterial microsymbiont of *Datisca glomerata*, expresses the
959 canonical *nod* Genes *nodABC* in symbiosis with its host plant. *PLoS One*, 10, e0127630.
960 Pérez Carrascal, O.M., VanInsberghe, D., Juárez, S., Polz, M.F., Vinuesa, P., and González, V. (2016).
961 Population genomics of the symbiotic plasmids of sympatric nitrogen-fixing *Rhizobium* species associated
962 with *Phaseolus vulgaris*. *Environ Microbiol*, 18, 2660-2676.
963 Pérez-Mendoza, D., Sepúlveda, E., Pando, V., Muñoz, S., Nogales, J., Olivares, J., et al. (2005).
964 Identification of the *rctA* gene, which is required for repression of conjugative transfer of rhizobial
965 symbiotic megaplasmids. *J Bacteriol*, 187, 7341-7350.

966 Ramsay, J.P., Sullivan, J.T., Jambari, N., Ortori, C.A., Heeb, S., Williams, P., et al. (2009). A LuxRI-family
967 regulatory system controls excision and transfer of the *Mesorhizobium loti* strain R7A symbiosis island by
968 activating expression of two conserved hypothetical genes. *Mol Microbiol*, 73, 1141-1155.
969 Ramsay, J.P., Tester, L.G.L., Major, A.S., Sullivan, J.T., Edgar, C.D., Kleffmann, T., et al. (2015). Ribosomal
970 frameshifting and dual-target antiactivation restrict quorum-sensing-activated transfer of a mobile
971 genetic element. *Proc Natl Acad Sci U S A*, 112, 4104-4109.
972 Remigi, P., Capela, D., Clerissi, C., Tasse, L., Torchet, R., Bouchez, O., et al. (2014). Transient
973 hypermutagenesis accelerates the evolution of legume endosymbionts following horizontal gene
974 transfer. *PLoS Biol*, 12, e1001942.
975 Remigi, P., Zhu, J., Young, J.P.W., and Masson-Boivin, C. (2016). Symbiosis within symbiosis: evolving
976 nitrogen-fixing legume symbionts. *Trends Microbiol*, 24, 63-75.
977 Reuhs, B.L., Relić, B., Forsberg, L.S., Marie, C., Ojanen-Reuhs, T., Stephens, S.B., et al. (2005). Structural
978 characterization of a flavonoid-inducible *Pseudomonas aeruginosa* A-band-like O antigen of *Rhizobium*
979 sp. strain NGR234, required for the formation of nitrogen-fixing nodules. *J Bacteriol*, 187, 6479-6487.
980 Rodelas, B., Lithgow, J.K., Wisniewski-Dye, F., Hardman, A., Wilkinson, A., Economou, A., et al. (1999).
981 Analysis of quorum-sensing-dependent control of rhizosphere-expressed (*rhi*) genes in *Rhizobium*
982 *leguminosarum* bv. viciae. *J Bacteriol*, 181, 3816-3823.
983 Rolin, D.B., Pfeffer, P.E., Osman, S.F., Szwergold, B.S., Kappler, F., and Benesi, A.J. (1992). Structural
984 studies of a phosphocholine substituted beta-(1,3);(1,6) macrocyclic glucan from *Bradyrhizobium*
985 *japonicum* USDA 110. *Biochim Biophys Acta*, 1116, 215-225.
986 Rosemeyer, V., Michiels, J., Verreth, C., and Vanderleyden, J. (1998). *luxI*- and *luxR*-homologous genes of
987 *Rhizobium etli* CNPAF512 contribute to synthesis of autoinducer molecules and nodulation of *Phaseolus*
988 *vulgaris*. *J Bacteriol*, 180, 815-821.

989 Saad, M.M., Crèvecoeur, M., Masson-Boivin, C., and Perret, X. (2012). The T3SS of *Cupriavidus*
990 *taiwanensis* strain LMG19424 compromizes symbiosis with *Leucaena leucocephala*. *Appl Environ*
991 *Microbiol,* 78, 7476-7479.
992 Schlaman, H.R., Okker, R.J., and Lugtenberg, B.J. (1992). Regulation of nodulation gene expression by
993 NodD in rhizobia. *J Bacteriol*, 174, 5177-5182.
994 Schwedock, J.S., and Long, S.R. (1992). *Rhizobium meliloti* genes involved in sulfate activation: the two
995 copies of *nodPQ* and a new locus, *saa*. *Genetics*, 132, 899-909.
996 Schäper, S., Steinchen, W., Krol, E., Altegoer, F., Skotnicka, D., Søgaard-Andersen, L., et al. (2017). AraC-
997 like transcriptional activator CuxR binds c-di-GMP by a PilZ-like mechanism to regulate extracellular
998 polysaccharide production. *Proc Natl Acad Sci U S A*, 114, E4822-E4831.
999 Sciotti, M.A., Chanfon, A., Hennecke, H., and Fischer, H.M. (2003). Disparate oxygen responsiveness of
1000 two regulatory cascades that control expression of symbiotic genes in *Bradyrhizobium japonicum*. *J*
1001 *Bacteriol*, 185, 5639-5642.
1002 Serrato, R.V. (2014). Lipopolysaccharides in diazotrophic bacteria. *Front Cell Infect Microbiol,* 4, 119.
1003 Skorupska, A., Janczarek, M., Marczak, M., Mazur, A., and Krol, J. (2006). Rhizobial exopolysaccharides:
1004 genetic control and symbiotic functions. *Microb Cell Fact*, 5, 7.
1005 Sugawara, M., Epstein, B., Badgley, B.D., Unno, T., Xu, L., Reese, J., et al. (2013). Comparative genomics
1006 of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. *Genome*
1007 *Biol*, 14, R17.
1008 Sugawara, M., Takahashi, S., Umehara, Y., Iwano, H., Tsurumaru, H., Odake, H., et al. (2018). Variation in
1009 bradyrhizobial NopP effector determines symbiotic incompatibility with Rj2-soybeans via effector-
1010 triggered immunity. *Nat Commun*, 9, 3139.
1011 Sullivan, J., and Ronson, C. (1998). Evolution of rhizobia by acquisition of a 500-kb symbiosis island that
1012 integrates into a phe-tRNA gene. *Proc Natl Acad Sci U S A*, 95, 5145-5149.

1013 Sullivan, J.T., Patrick, H.N., Lowther, W.L., Scott, D.B., and Ronson, C.W. (1995). Nodulating strains of
1014 *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. *Proc Natl Acad*
1015 *Sci U S A*, 92, 8985-8989.
1016 Svistoonoff, S., Benabdoun, F.M., Nambiar-Veetil, M., Imanishi, L., Vaissayre, V., Cesari, S., et al. (2013).
1017 The independent acquisition of plant root nitrogen-fixing symbiosis in Fabids recruited the same genetic
1018 pathway for nodule organogenesis. *PLoS One*, 8, e64515.
1019 Sánchez, C., Iannino, F., Deakin, W.J., Ugalde, R.A., and Lepek, V.C. (2009). Characterization of the
1020 *Mesorhizobium loti* MAFF303099 type-three protein secretion system. *Mol Plant Microbe Interact*, 22,
1021 519-528.
1022 Tang, F., Yang, S., and Zhu, H. (2016). Functional analysis of alternative transcripts of the soybean Rj2
1023 gene that restricts nodulation with specific rhizobial strains. *Plant Biol (Stuttg)*, 18, 537-541.
1024 Tian, C.F., Zhou, Y.J., Zhang, Y.M., Li, Q.Q., Zhang, Y.Z., Li, D.F., et al. (2012). Comparative genomics of
1025 rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations.
1026 *Proc Natl Acad Sci U S A*, 109, 8629-8634.
1027 Tun-Garrido, C., Bustos, P., González, V., and Brom, S. (2003). Conjugative transfer of p42a from
1028 *Rhizobium etli* CFN42, which is required for mobilization of the symbiotic plasmid, is regulated by
1029 quorum sensing. *J Bacteriol*, 185, 1681-1692.
1030 Valls, M., Genin, S., and Boucher, C. (2006). Integrated regulation of the type III secretion system and
1031 other virulence determinants in *Ralstonia solanacearum*. *PLoS Pathog*, 2, 798-807.
1032 Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., et al. (2010). Plant
1033 peptides govern terminal differentiation of bacteria in symbiosis. *Science*, 327, 1122-1126.
1034 van Workum, W.A.T., van Slageren, S., van Brussel, A.A.N., and Kijne, J.W. (1998). Role of
1035 exopolysaccharides of *Rhizobium leguminosarum* bv. *viciae* as host plant-specific molecules required for
1036 infection thread formation during nodulation of *Vicia sativa. Mol Plant Microbe Interact*, 11, 1233–1241.

1037 Vernié, T., Kim, J., Frances, L., Ding, Y., Sun, J., Guan, D., et al. (2015). The NIN transcription factor
1038 coordinates diverse nodulation programs in different tissues of the *Medicago truncatula* root. *Plant Cell*,
1039 27, 3410-3424.
1040 Viprey, V., Del Greco, A., Golinowski, W., Broughton, W.J., and Perret, X. (1998). Symbiotic implications
1041 of type III protein secretion machinery in *Rhizobium*. *Mol Microbiol*, 28, 1381-1389.
1042 Wang, D., Couderc, F., Tian, C.F., Gu, W., Liu, L.X., and Poinsot, V. (2018). Conserved composition of Nod
1043 factors and exopolysaccharides produced by different phylogenetic lineage. *Front Microbiol*, 9, 2852.
1044 Wippel, K., and Long, S.R. (2018). Symbiotic performance of *Sinorhizobium meliloti* lacking ppGpp
1045 depends on the *Medicago* host species. *Mol Plant Microbe Interact*, doi: 10.1094/MPMI-11-18-0306-R.
1046 Xin, D.W., Liao, S., Xie, Z.P., Hann, D.R., Steinle, L., Boller, T., et al. (2012). Functional analysis of NopM, a
1047 novel E3 ubiquitin ligase (NEL) domain effector of *Rhizobium* sp. strain NGR234. *PLoS Pathog*, 8,
1048 e1002707.
1049 Yang, M., Sun, K., Zhou, L., Yang, R., Zhong, Z., and Zhu, J. (2009). Functional analysis of three AHL
1050 autoinducer synthase genes in *Mesorhizobium loti* reveals the important role of quorum sensing in
1051 symbiotic nodulation. *Can J Microbiol*, 55, 210-214.
1052 Yang, S.M., Tang, F., Gao, M.Q., Krishnan, H.B., and Zhu, H.Y. (2010). R gene-controlled host specificity in
1053 the legume-rhizobia symbiosis. *Proc Natl Acad Sci U S A*, 107, 18735-18740.
1054 Zhang, L., Chen, X.J., Lu, H.B., Xie, Z.P., and Staehelin, C. (2011). Functional analysis of the type 3 effector
1055 nodulation outer protein L (NopL) from *Rhizobium* sp. NGR234: symbiotic effects, phosphorylation, and
1056 interference with mitogen-activated protein kinase signaling. *J Biol Chem*, 286, 32178-32187.
1057 Zhao, R., Liu, L.X., Zhang, Y.Z., Jiao, J., Cui, W.J., Zhang, B., et al. (2018). Adaptive evolution of rhizobial
1058 symbiotic compatibility mediated by co-evolved insertion sequences. *ISME J*, 12, 101-111.

1059 Zheng, H., Zhong, Z., Lai, X., Chen, W.X., Li, S., and Zhu, J. (2006). A LuxR/LuxI-type quorum-sensing
1060 system in a plant bacterium, *Mesorhizobium tianshanense*, controls symbiotic nodulation. *J Bacteriol*,
1061 188, 1943-1949.
1062 Österman, J., Marsh, J., Laine, P.K., Zeng, Z., Alatalo, E., Sullivan, J.T., et al. (2014). Genome sequencing of
1063 two *Neorhizobium galegae* strains reveals a *noeT* gene responsible for the unusual acetylation of the
1064 nodulation factors. *BMC Genomics*, 15, 500.
1065

1066 **List of tables**

1067 **Table 1**. **Bacterium genera containing rhizobia and examples of rhizobium species.**
1068 * Rhizobium taxonomy is based on (Estrada-de Los Santos et al., 2018; Mousavi et al., 2015; Ormeño-
1069 Orrillo et al., 2015).

1070

1071 **Table 2**. **Presence of some NF biosynthesis and export, EPS biosynthesis and nitrogen fixation genes in**
1072 **various rhizobia.**
1073 Grey boxes mark the presence of the gene. Ac, *Azorhizobium caulinodans* ORS571. Bd, *Bradyrhizobium*
1074 *diazoefficiens* USDA110. Bs, *Bradyrhizobium* sp. ORS278. Ml, *Mesorhizobium loti* MAF303099. Re,
1075 *Rhizobium etli* CFN42. Rl, *Rhizobium leguminosarum* bv. viciae 3841. Sm, *Sinorhizobium meliloti* 1021. Sf,
1076 *Sinorhizobium fredii* NGR234. Ct, *Cupriavidus taiwanensis* LMG19424. Pp, *Paraburkholderia phymatum*
1077 STM815. * Regulatory genes are in boldface. # (Daubech et al., unpublished). † *nifT* is also known as *fixU*
1078 or *fixT*. Adapted and updated from Masson-Boivin et al., 2009 and Black et al., 2012 and completed
1079 based on gene annotations from MaGe
1080 (http://www.genoscope.cns.fr/agc/microscope/home/index.php).
1081

1082 **Figure legend**

1083 **Figure 1: Rhizobial genetic factors used in nodulation, infection and nitrogen fixation.**

1084 **A.** In the rhizosphere, plants secrete flavonoids perceived by NodDs proteins which induce the
1085 expression of *nod* genes. *nod* gene expression leads to the production of Nod Factors specifically
1086 recognized by plant receptors and induces the symbiotic signaling pathway (CSSP). In the *nod*-
1087 independent rhizobia, the mechanism inducing the CSSP is not known. During the infection process,
1088 rhizobia also synthesize a panel of surface polysaccharides (EPS, LPS, CGs) to evade plant defense
1089 responses. In certain rhizobia, establishment of symbiosis is assisted by T3 or T4 secretion systems unless
1090 secreted effectors are recognized by plant resistance proteins.

1091 **B.** In nodule cells, released bacteria from infection threads remained surrounded by a peribacteroid
1092 membrane of plant origin thus forming symbiosomes in the plant cytosol. The peribacteroid space is
1093 probably an acidic compartment. LPS are required for the survival of bacteria following their release in
1094 nodule cells. In symbiosomes, bacteria differentiate into bacteroids able to fix atmospheric nitrogen into
1095 ammonia. Bacteroid terminal differentiation triggered by NCR peptides occurred in plants from IRLC and
1096 Dalbergioid clades and necessitates the bacterial BacA peptide transporter. Nitrogen fixation is possible
1097 in the microoxic environment of symbiosomes. A high oxygen affinity cytochrome *cbb*3 oxidase allowed
1098 the microaerobic respiration of bacteroids. Carbon sources are provided by the plant as C_4 dicarboxylic
1099 acids.

1100 The symbiotic determinants illustrated in panels A and B are from different bacteria. NFs, Nod Factors.
1101 QS, quorum sensing molecules. T3SS, type III secretion system. EPS, exopolysaccharides. LPS,
1102 lipopolysaccharides. CGs, cyclic β-glucans. NFRs, Nod factor receptors. EPRs, EPS receptors. CSSP,
1103 common symbiotic signaling pathway. R, resistance proteins. TCA, tricarboxylic acid cycle.

1104

1105 **Figure 2: Experimental evolution of legume symbionts**

1106 **A. Proposed evolutionary model for the emergence of NF-dependent rhizobia** (adapted from Masson-
1107 Boivin and Sachs, 2018).
1108 Symbiotic plasmids or islands carrying essential symbiotic genes are horizontally transferred in the soil
1109 from rhizobia to non-rhizobium strains. These transfers may lead to a proto-rhizobium that needs
1110 additional genome remodeling under the plant selection pressure to turn into an effective nitrogen fixing
1111 rhizobium.
1112 **B. Experimental evolution of the plant pathogen *R. solanacearum* into legume symbionts** (Masson-
1113 Boivin and Sachs, 2018)**.**
1114 The *nod-nif* symbiotic genes from *Cupriavidus taiwanensis*, the symbiont of *Mimosa pudica*, were
1115 introduced into the plant pathogen *Ralstonia solanacearum*. The resulting chimeric strain was able to
1116 produce Nod Factors and induce root hair curling but unable to form nodules on *M. pudica*. The ability to
1117 perform the two first steps of symbiosis, nodulation and intracellular infection, was progressively
1118 acquired in parallel lineages evolved by repeated inoculation to *M. pudica* (Marchetti et al., 2010;
1119 Marchetti et al., 2017). Mutations allowing the main symbiotic shifts were identified (Capela et al., 2017;
1120 Guan et al., 2013; Marchetti et al., 2010). Nodulation was acquired through inactivation of the T3SS
1121 either via a *hrcV* or a *hrpG* stop mutation. A first level of intracellular infection was acquired via a *hrpG*
1122 stop mutation or a combination of a *hrcV* mutation and inactivation of other virulence regulators such as
1123 *hrpG*, *prhI* or *vsrA*. Optimization of intracellular infection was observed following mutations affecting
1124 global virulence regulatory pathways such as *efpR* or the *phc* quorum-sensing pathway.
1125 **C. Rewiring of the *R. solanacearum* virulence regulatory network leading to symbiotic adaptation.**
1126 The virulence master regulator HrpG not only controls the T3SS via HrpB but also *ca.* 200 other genes
1127 some of them interfere with intracellular infection of *M. pudica* nodules. The Phc quorum sensing system
1128 controls the activity of PhcA, which in turn regulates the expression of hundreds of genes including
1129 components of the HrpG regulatory cascade, EPS production, motility and cell wall degrading enzymes.

1130 VsrA and EfpR are two other virulence global regulators controlling similar functions as PhcA. Genes

1131 mutated in the evolution experiment are colored according to their associated symbiotic phenotypes.

1132 QS, quorum sensing molecules. EPS, exopolysaccharides. T3SS, type 3 secretion system. T3E, type 3

1133 effectors. Dashed and straight lines indicate indirect and direct controls, respectively.

1134

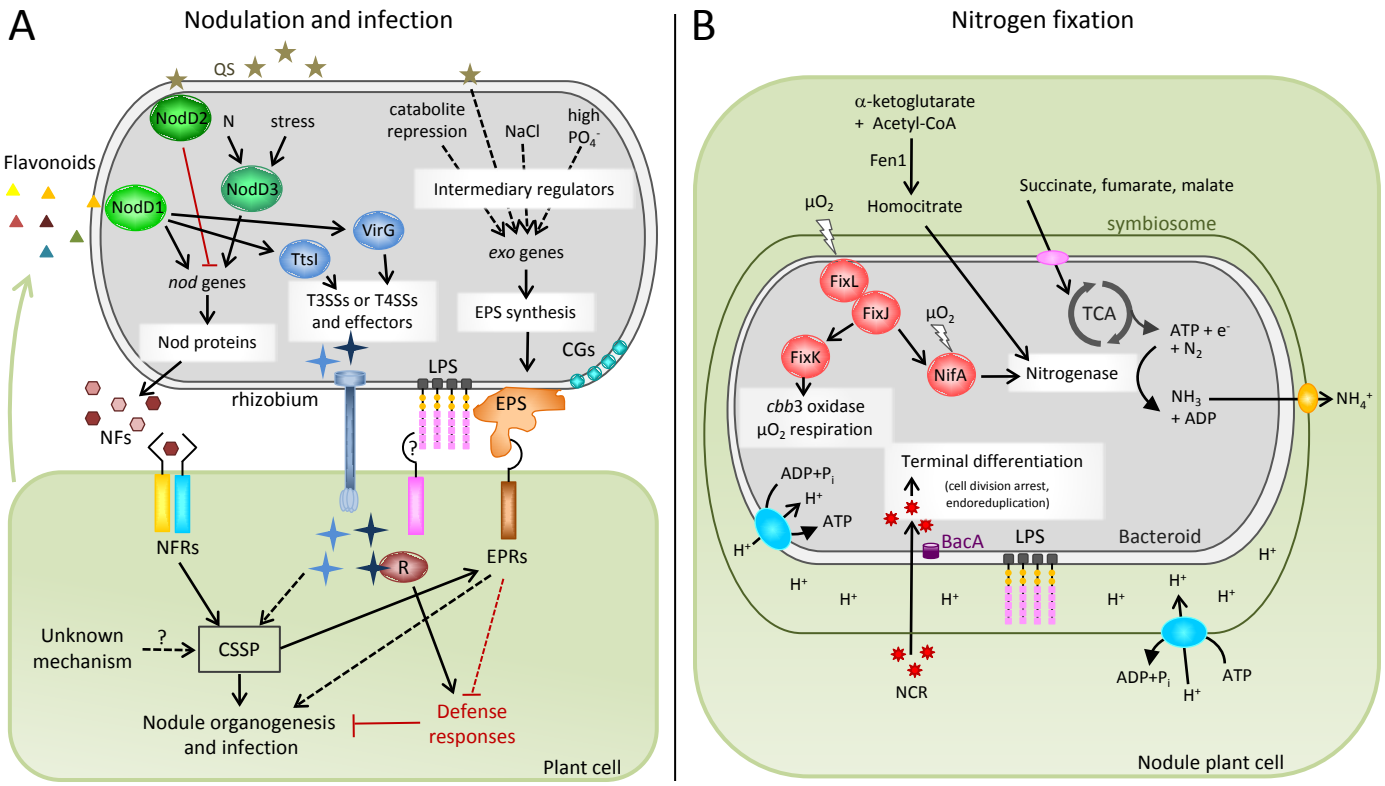
A
Nodulation and infection
QS
NodD2
N
stress
catabolite repression
NaCl
high PO_4^-
Flavonoids
NodD3
NodD1
Intermediary regulators
VirG
TtsI
exo genes
nod genes
T3SSs or T4SSs and effectors
EPS synthesis
Nod proteins
CGs
LPS
rhizobium
EPS
NFs
?
NFRs
EPRs
R
Unknown mechanism
?
CSSP
Nodule organogenesis and infection
Defense responses
Plant cell
B
Nitrogen fixation
α-ketoglutarate + Acetyl-CoA
Fen1
Homocitrate
Succinate, fumarate, malate
μO_2
symbiosome
FixL
FixJ
μO_2
TCA
ATP + e^- + N_2
FixK
NifA
Nitrogenase
cbb3 oxidase μO_2 respiration
NH_3 + ADP
NH_4^+
Terminal differentiation
(cell division arrest, endoreduplication)
ADP+P_i
H^+
ATP
BacA
LPS
Bacteroid
NCR
ADP+P_i
ATP
Nodule plant cell

Figure 1

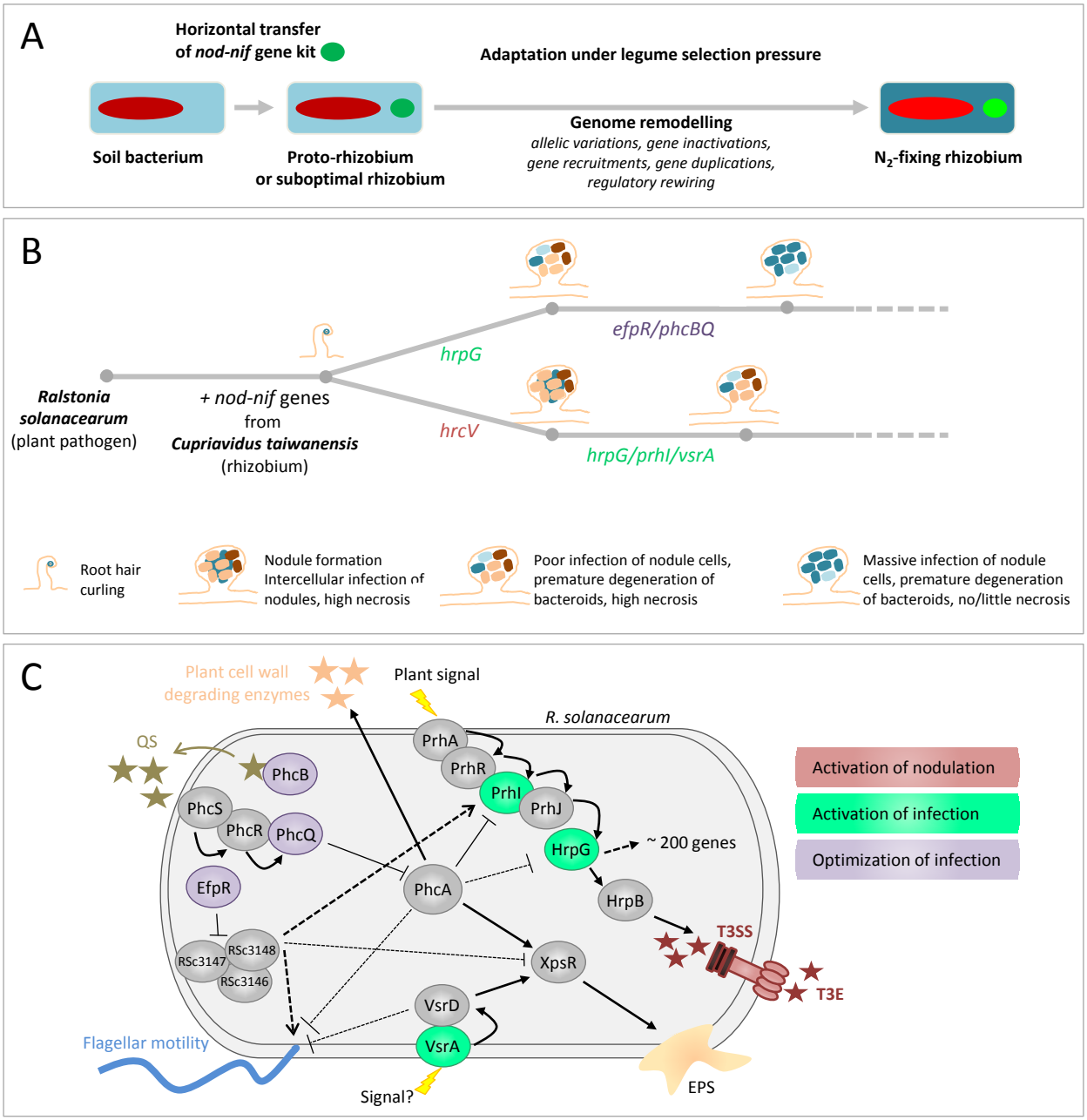
A
Horizontal transfer
of nod-nif gene kit
Adaptation under legume selection pressure
Soil bacterium
Proto-rhizobium
or suboptimal rhizobium
Genome remodelling
allelic variations, gene inactivations,
gene recruitments, gene duplications,
regulatory rewiring
N2-fixing rhizobium
B
Ralstonia
solanacearum
(plant pathogen)
+ nod-nif genes
from
Cupriavidus taiwanensis
(rhizobium)
hrpG
hrcV
efpR/phcBQ
hrpG/prhI/vsrA
Root hair
curling
Nodule formation
Intercellular infection of
nodules, high necrosis
Poor infection of nodule cells,
premature degeneration of
bacteroids, high necrosis
Massive infection of nodule
cells, premature degeneration
of bacteroids, no/little necrosis
C
Plant cell wall
degrading enzymes
Plant signal
R. solanacearum
QS
PhcB
PhcS
PhcR
PhcQ
EfpR
RSc3148
RSc3147
RSc3146
PhcA
PrhA
PrhR
PrhI
PrhJ
HrpG
~ 200 genes
HrpB
XpsR
VsrD
VsrA
T3SS
T3E
EPS
Flagellar motility
Signal?
Activation of nodulation
Activation of infection
Optimization of infection

Figure 2

Table 1: Bacterium genera containing rhizobia and examples of rhizobium species.

Class of proteobacteria	Genera*	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*

* Rhizobium taxonomy is based on (Mousavi et al., 2015 ; Ormeno-Orrillo et al., 2015 ; Estrada-de Los Santos et al., 2018).

Table 2. Presence of some NF biosynthesis and export, EPS biosynthesis and nitrogen fixation genes in various rhizobia.

Functions	Genes*	Ac	Bd	Bs	Ml	Re	Rl	Sm	Sf	Ct	Pp
Nod factor synthesis and export genes	*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		■		■				■		
	nolT				■						
	nolU		■		■				■		
	nolV		■		■						
	nolW				■						
	nolX				■						
EPS biosynthesis genes	*exoA*			■	■			■	■		
	exoB	■	■	■	■	■	■	■	■		
	exoD	■				■	■	■	■		
	exoF				■	■	■	■	■		
	exoH								■		
	exoI			■	■	■	■	■	■		
	exoK, exoL				■			■	■		
	exoM, exoQ				■	■	■	■	■		
	exoN	■	■	■	■	■	■	■	■		
	exoO				■			■	■		
	exoP		■	■	■	■	■	■	■		
	exoR, exoS, chvI	■	■	■	■	■	■	■	■		
	exoT					■	■				
	exoU				■	■	■	■	■		
	exoV					■	■	■	■		
	exoX						■				
	exoY	■	■	■	■	■	■	■	■		
	exoZ				■	■	■	■	■		
	pssA	■	■	■	■	■	■	■	■	■	
	pssB	■	■	■	■	■	■	■	■		
	pssN	■	■	■	■	■	■	■	■		
	pssO					■	■				
	pssP		■	■	■	■	■	■	■		
	pssT		■	■		■	■				
	pssVSRMKLJIHGFCDE					■	■				
Nitrogen fixation genes	***fixL***	■	■	■	■	■	■	■			
	fixJ	■	■	■	■	■		■			
	fixK	■	■	■	■	■	■	■	■		
	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. * Regulatory genes are in boldface. # (Daubech et al., unpublished data). † *nifT* is also known as *fixU* or *fixT*. Adapted and updated from Masson-Boivin et al., 2009 and Black et al., 2012 and completed based on gene annotations from MaGe (http://www.genoscope.cns.fr/agc/microscope/home/index.php).