DOTTORATO DI RICERCA IN
Scienze Agrarie ed Ambientali

CICLO XXXIII

COORDINATORE Prof. Giacomo Pietramellara

Into the wild: how rhizobia compete and survive in the early stage of symbiosis

Settore Scientifico Disciplinare AGR/16

Dottorando
Dott.ssa Bellabarba Agnese

Tutore
Prof. Viti Carlo

Coordinatore
Prof. Pietramellara giacomo

Co-tutore
Prof. Alessio Mengoni

Anni 2017/2020

Summary

Associations between leguminous plants and symbiotic nitrogen-fixing bacteria (rhizobia) are a classic example of mutualism between a eukaryotic host and a specific group of prokaryotic microbes. Rhizobia improve plant yield furnishing fixed nitrogen; therefore, they are highly used as inoculants, especially in sustainable agriculture. Though this association is species-specific, legume roots are exposed to heterogeneous rhizobial populations where different compatible strains are present and, could be infected by more than one strain.

It is known that within the same rhizobial species different strains may have different competition capabilities, but detailed analyses able to predict the rhizobial competitive phenotype based exclusively on their genome are still lacking. In this thesis, we performed a bacterial genome-wide association analysis (GWAS) to define which genetic traits are responsible for an improved competitive phenotype in the model species *Sinorhizobium meliloti*. A panel of thirteen *S. meliloti* strains, whose genome is completely sequenced, has been selected and tested against three *S. meliloti* reference competitor strains (Rm1021, AK83 and BL225C) in a *Medicago sativa* nodule occupancy test. The measure of competition phenotypes previously obtained in the competition tests, in combination with strains genomic sequences of strains tested, were used to build-up a k-mer-based statistical models for each set of competition experiments. The obtained models were then applied to evaluate the accuracy in predicting the competition abilities of strains in the three competition patterns (vsRm1021, vsAK83 and vsBL225C). The competitive abilities of *S. meliloti* strains against two partners, BL225C and Rm1021, were well predict by the predictive models, as shown by the coefficient of determination R^2 (equal to 0.96 and 0.84, respectively). Four strains showing the highest competition phenotypes (> 60% single strain nodule occupancy; GR4, KH35c, KH46 and SM11) versus BL225C were used to identify *k-mers* associated with the competition phenotype. The most significantly associated *k-mers* ($p < 0.05$) were mapped on the genomic sequences of the *S. meliloti* strains used. Most of the *k-mers* were located on the symbiosis-related megaplasmid pSymA and on genes coding for transporters, proteins involved in the biosynthesis of cofactors and proteins related to metabolism (i.e. glycerol, fatty acids) suggesting that competition abilities reside in multiple genetic determinants comprising several cellular components.

The identification of the best rhizobial inoculants in two new effective breeding lines of pea (*Pisum sativum*) used in Lithuania (DS 3637-2 and DS 3795-3) was also investigated. Six

rhizobial strains, isolated from pea plants, which could be used as potential inoculants, were phylogenetically identified and extensive phenotypically characterized by Phenotype Microarray. All the strains belonged to the *Rhizobium leguminosarum* group, and were subdivided into three groups related to *Rhizobium anhuiense*, *Rhizobium leguminosarum* bv. *viciae* and *R. sophorae/R. laguerreae*. Differences observed with Phenotype microarray were linked to different phylogeny of the strains. In term of symbiotic efficiency, six strains showed different symbiotic performances depending on the breeding line used. In particular, *Rhizobium anhuiense* strain Z1 (the reference strain) and *Rhizobium leguminosarum* bv. *viciae* 14ZE were the best symbiotic inoculants with breeding lines DS 3637-2 and DS 3795-3, respectively.

Summary

Chapter 1 – Introduction

1.1 Socio-microbiology of legume-rhizobia symbioses: interaction strategies.. pp. 2

1.1.1 The legume - rhizobia symbioses……………………………………… pp. 2

1.1.1.1 First steps: nodulation signaling and attachment to roots………. pp. 5

1.1.1.2 Infection, nodule organogenesis and type of nodules………….. pp. 8

1.1.1.3 The nitrogen fixation process and nutrients exchanges……….. pp. 12

1.1.2 Infecting legumes or forgoing symbiosis? ………………………… pp. 16

1.1.3 Cooperation and conflicts of interest……………………………… pp. 19

1.1.4 Features of ineffective strains……………………………………... pp. 20

1.1.5 The "Tragedy of the Commons" and a new concept of cheating…. pp. 22

1.1.6 Host sanctions and other stabilizing evolutionary mechanisms…... pp. 24

1.1.7 Why do cheater rhizobia still exist?……………………………..… pp.27

1.1.8 Single-strain inoculation *versus* co-inoculation..………………… pp. 30

1.2 The ecology of competition……………………………………………… pp. 33

1.2.1 Features that affect nodulation competitiveness in rhizobia………………. pp. 34

1.2.2 Methodological approaches to study the competition for nodulation…….. pp. 36

1.2.3 The competition problem in agricultural settings………………………… pp. 38

1.3 The quest for elite inoculant strains: integrated genomics approaches……………………………………………………………………... pp. 38

1.4 References……………………………………………………………….. pp. 41

Chapter 2 - Aim of the thesis…………………………………………….. pp. 60

Chapter 3 - Competitiveness prediction for nodule colonization in *Sinorhizobium meliloti* through combined *in vitro* tagged strain characterization and genome-wide association analysis………………………………………………………… pp. 62

Chapter 4 - Deciphering the symbiotic plant microbiome: traslating the most recent discoveries on rhizobia for the improvement of agricultural practices in metal-contaminated and high saline lands……………………………………………... pp. 114

Chapter 5 - Selection of *Rhizobium* strains for inoculation of Lithuanian breeding lines of *Pisum sativum*…………………………………………………….. pp. 155

Chapter 6 - Concluding remarks………………………………………… pp. 201

Chapter 1

Introduction

1.1 Socio-microbiology of legume-rhizobia symbioses: interaction strategies

1.1.1 The legume - rhizobia symbioses

The exclusively prokaryotic process of biological nitrogen fixation (BNF), by which atmospheric nitrogen (N_2 chemically inert) is enzymatically reduced to the metabolically accessible form ammonia (NH_3), represents the largest input of nitrogen to the biosphere underpinning the global nitrogen cycle (Lodwig and Poole, 2003; Peters et al., 2006; Gibson et al., 2008). The BNF occurs in different bacteria, including in rhizobia, cyanobacteria, azobacteria, *Frankia*, and some strains of *Archaea* (Lodwig and Poole, 2003; Raymond et al., 2004; Gibson et al., 2008). Neverteless, the greatest contribution to the BNF process comes from the symbiotic association between rhizobia and plants of the *Leguminosae* family (or *Fabaceae*) (Lodwig and Poole, 2003). Indeed, rhizobia-legume symbioses provide more than half of the world's biologically fixed nitrogen (Smil et al., 2001), an amount approximately equal to 40–48 million tonnes of nitrogen introduced into agricultural systems each year (Herridge et al., 2008). Therefore, the economic relevance of this association is enormous: precisely, the impact of legume-rhizobia symbiosis on the global agricultural economy was estimated to be worth the equivalent of USD160–180 billion (Rajwar et al., 2013). For more than 100 years, BNF has captured the attention of scientists, and it has been extensively exploited in sustainable agriculture (Dixon and Wheeler, 1986). Indeed, the study of the symbiosis between rhizobia and plants is one of the greatest contributions of microbiology to agricultural applications, designed to improve the yield of leguminous crops and their cultivation as fodder plants, with the possibility to recovering degraded areas (Bellabarba et al., 2019).

Rhizobia are a polyphyletic group of Gram-negative soil bacteria composed of the *Rhizobiaceae* family of the Alpha-proteobacteria subdivision and some genera of Beta-proteobacteria (*Cupriavidus* and *Burkholderia*). In particular, the family *Rhizobiaceae* comprise several genera, including *Rhizobium*, *Sinorhizobium*, *Mesorhizobium*, *Bradyrhizobium*, and *Azorhizobium*. Therefore. the general term "rhizobia" is referred to a genetically diverse and physiologically heterogeneous group of N_2-fixing soil bacteria able to infect and colonize plant roots establishing a mutualistic symbiosis with leguminous plants. Their capability to improve

plant growth is primarily based on their ability to fix nitrogen which is provided to the plant (Wheatley and Poole, 2018). After attaching to and colonizing legume roots, rhizobia become internally housed in protective 'nodule' structures. Within nodules rhizobia differentiate into bacteroids, the N_2-fixing form of rhizobia (Poole et al., 2018). Bacteroids are surrounded by the bacterial and plant-derived membranes forming the symbiosome, which can be defined as the engine of fixation (Poole et al., 2018). Low O_2 concentrations in the symbiosome, and in general in the nodule, allow nitrogenase to reduce N_2 to ammonia (Gage, 2004; Gibson et al., 2008). In this symbiosis, rhizobia provide the plant with useable nitrogen sources for growth (Udvardi and Poole, 2013) which the plant uses to synthesize amino acids (Poole et al., 2018). In return to the exported ammonia, rhizobia receive carbon sources primarily in the form of dicarboxylates, whose metabolism helps to fuel the energetically expensive process of nitrogen fixation (Udvardi and Poole, 2013).

There is often a host-symbiont specificity in the establishment of a nitrogen-fixing symbiosis (Poole et al., 2018). For example, *Rhizobium leguminosarum* bv. *trifolii* and *Sinorhizobium meliloti* infect only *Trifollium* and *Medicago* species, respectively (Poole et al., 2018). In contrast to these restricted host range rhizobia, few rhizobia are promiscuous, as *Sinorhizobium fredii* NGR234 which infects 112 different legume genera (Relic et al., 1993; Pueppke and Broughton, 1999). Based on rhizobial phylogenetic relationships, it was suggested that the restricted host range symbiosis evolved from an ancestral broad host range symbiosis (Gibson et al., 2008), creating a finely tuned symbiosis. The most representative rhizobial species and their plant host species are listed in Table 1.

Table 1. Rhizobia and their legume hosts. *This table lists representative rhizobial species and strains; a comprehensive list of most of the known rhizobia and their plant hosts has recently been compiled in Andrews et al., 2017 and Sprent et al., 2017. From Poole et al. (2018), modified.

Rhizobia*	Plant host species
Azorhizobium caulinodans	*Sesbania rostrate*
Azorhizobium doebereinerae	*Sesbania virgate*
Bradyrhizobium japonicum	*Glycine max*
Bradyrhizobium spp.	*Cajanus cajun, Glycine max, Aeschynomene americana, Aeschynomene indica, Aeschynomene rudis, Arachis duranensis* and *Arachis hypogaea*
Burkholderia spp.	*Mimosa pudica*
Cupriavidus sp. AMP6	*Mimosa asperata*
Cupriavidus taiwanensis	*Mimosa pudica*
Sinorhizobium meliloti	*Medicago sativa* and *Medicago truncatula*
Mesorhizobium haukuii	*Sesbania sesban*
Mesorhizobium loti	*Lotus japonicus* and *Lotus corniculatus*
Rhizobium etli	*Phaseolus vulgaris*
Rhizobium gallicum	*Phaseolus vulgaris*
Rhizobium leguminosarum bv. *phaseoli*	*Phaseolus vulgaris*
Rhizobium leguminosarum bv. *trifolii*	*Trifollium* spp.
Rhizobium leguminosarum bv. *viciae*	*Pisum sativum, Lens culinaris, Vicia cracca, Vicia hirsuta* and *Vicia faba*
Rhizobium sp. strain IRBG74	*Sesbania cannabina, Sesbania bispinosa, Sesbania cannabina, Sesbania exasperata, Sesbania formosa, Sesbania grandiflora, Sesbania madagascariensis, Sesbania macrantha* and *Sesbania pachycarpa*
Rhizobium tropici	*Phaseolus vulgaris*
Sinorhizobium fredii NGR234	112 genera
Sinorhizobium fredii USDA257	79 genera
Sinorhizobium medicae	*Medicago sativa* and *Medicago truncatula*

1.1.1.1 First steps: nodulation signaling and attachment to roots

The mechanism which allows the symbiotic interaction between the nitrogen-fixing rhizobia and leguminous plants is under a complex network of molecular (genetic and metabolic) interactions. Indeed, both organisms have evolved a complex and specific signals exchange that schedules this multi-step mutualist association (Jones et al., 2017). The mutual recognition between partners occurs in the rhizosphere, which is the soil area that immediately surrounding plant roots characterized by a large variety of microorganisms. The rhizosphere represents an important source of nutrients for bacteria due to the presence of root exudates (Gibson et al., 2008). Indeed, in this area, aromatic amino acids are greatly mobilized to fuel bacterial metabolism (Ramachandran et al., 2001). Rhizobia are then attracted towards the rhizosphere by chemoattractants present in root exudates (Poole et al., 2018).

The symbiotic process starts with the perception by rhizobia of flavonoid compounds (as luteolin) released from the plant roots (Gibson et al., 2008). Indeed, plant-derived flavonoids trigger a significant transcriptional response in the compatible rhizobia, mediated by the transcriptional regulator nodulation process NodD, which results in the production and the secretion of Nodulation Factors (NFs) (lipo-chito-oligosaccharide molecules, LCOs) (Capela et al., 2005) (Figure 1). The Nod factors, in turn, activate the symbiosis (SYM) signaling pathway on plant root cells. These signaling controls flavonoid synthesis creating a positive feedback loop between LCOs and flavonoids, which leads to the root hair curling, to the rhizobia infection, and ultimately to the nodule formation (Poole et al., 2018) (Figure 1). The exact chemical structures of the Nod factors and the flavonoids secreted by the plant are the basis for host-symbiont specificity. Common *nod* genes (*nodABC*), found in nearly all rhizobial species, are responsible for the synthesis of the shared core structure of LCOs, the chitin backbone (Poole et al., 2018). Indeed, mutations in these genes completely abolish the nodulation process and thus the symbiosis (Wang et al., 2012). In contrast, other *nod* genes confer specificity for nodulation of particular host plants and are involved in various modifications of the chitin backbone (Gibson et al., 2008).

The bacterial colonization of roots and nodulation signaling are interlinked processes. Indeed, the rhizobial attachment to roots hairs amplifies flavonoid – Nod Factors signaling (Pini et al., 2017) leading to root hair curling and entrapment of the rhizobia (Poole et al., 2017). Bacterial colonization of roots and attachment to root hairs has received less attention than nodulation signaling. Like other members of the rhizosphere community, the attachment to the plant roots

hairs by rhizobia occurs as a biphasic mechanism (Wheatley and Poole, 2018). This process is indeed characterized by a weak, reversible and non-specific primary attachment phase, followed by a stronger and more specific secondary attachment phase in which the synthesis of extracellular fibrils aids further bacterial accumulation and aggregation (Wheatley and Poole, 2018). In acidic soils, the rhizobial surface polysaccharides glucomannan binds to the plant lectin expressed on root hair tip (Laus et al., 2005) (Figure 1). Lectins are carbohydrate-binding proteins that operate as a receptor for rhizobial surface polysaccharides, binding them in a reversible and highly specific manner (Lagarda-Diaz et al., 2017; Wheatley and Poole, 2018). This lectin-mediated mechanism of primary attachment was firstly detected in the attachment of *B. japonicum* to soybean roots (Lodeiro et al. 2000; Wheatley and Poole, 2018) but is likely utilized in other *Rhizobium*–legume symbioses (Wheatley and Poole, 2018). Under alkaline and neutral conditions, the mechanism of a primary attachment may involve an extracellular calcium-binding rhizobial protein 'rhicadhesin' (Downie, 2010), though this dynamic remains poorly understood (Poole et al., 2018). Once the primary attachment to root hairs occurs, rhizobia may form a biofilm whose structure is called 'root hair cap' (Poole et al., 2018). This secondary attachment involved the *Rhizobium*-adhering proteins (Raps), extracellular polysaccharides (EPSs), and cellulose fibrils (Figure 1). The strong and direct binding of RapA2 to acidic and capsular EPS (Abdian et al., 2013) stabilizes biofilms, whereas cellulose fibrils are important for a tight adherence between rhizobial cells during biofilm formation (Poole et al., 2018) (Figure 1).

Lipopolysaccharides (LPS) may also be important for attachment to roots in *R. leguminosarum* bv *viciae* (Vanderline et al., 2009). Furthermore, a possible role of LPS in regulating the host immune response was hypothesized. LPSs are complex macromolecules composed of a lipid A membrane-anchor and an oligosaccharide core, which can be further modified by the addition of a variable O-antigen polysaccharide (Raetz and Whitfield, 2002). In rhizobia the unique structural modification of lipid A may contribute to symbiosis, shielding the rhizobial LPS from recognition by the plant's innate immune system (Gibson et al., 2008). Moreover, in *S. meliloti* the lipid A component of LPS is capable of suppressing the oxidative burst and, accordingly, any potential immune response to intracellular bacteria (Scheidle et al., 2005, Tellstrom et al., 2007). However, the precise function of LPS in promoting symbiosis remains unclear (Becker et al., 2005). Other mechanisms of modulation of the plant immune response were hypothesized (graphically reported in Figure 1) (Poole et al., 2018).

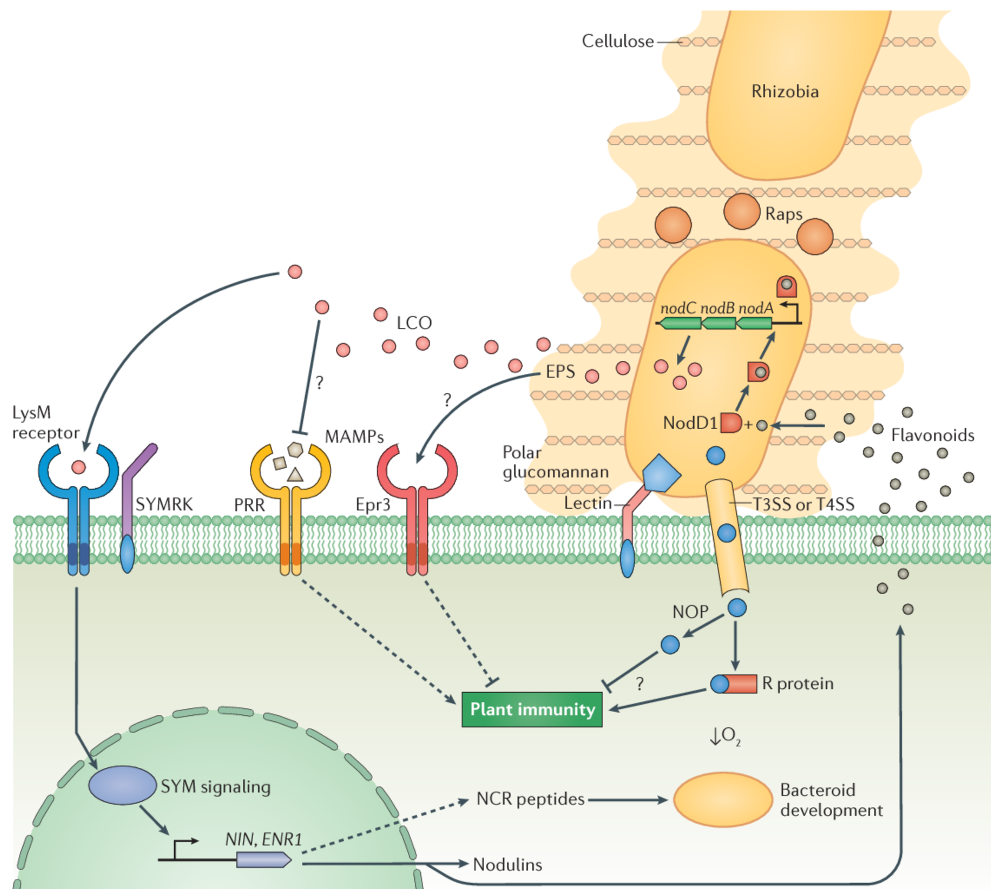


Figure 1. Molecular mechanisms of plant–rhizobia signaling. Many of the plant–rhizobia interactions take place on root hairs, where bacteria will form a root cap that is crosslinked by extracellular polysaccharides (EPS), *Rhizobium*-adhering proteins (Raps), and cellulose. Lipochitooligosaccharides (LCOs; also known as Nod factors) may bind the pattern recognition receptors (PRRs) to suppress the MAMP response (microorganism-associated molecular pattern). Similarly, the binding of EPS to the EPS receptor 3 (Epr3) suppresses immunity. Plant flavonoids induce the production of LCOs, which in turn bind a lysine motif (LysM) receptor heterocomplex. This activates the leucine-rich repeat protein receptor-like kinase (SYMRK) and, in turn, the symbiosis (SYM) signaling pathway. These signaling events result in the induction of nodule formation, including the production of nodule cysteine-rich (NCR) peptides, nodulins, and flavonoids. Some bacteria inject effector proteins (NOPs) through type III secretion systems (T3SSs) and type IV secretion systems (T4SSs). NOPs can increase the host range, presumably by suppressing plant immunity, although they may also bind to a plant R protein to stimulate immunity. NodD1, nodulation protein D1. From Poole et al. (2018), modified.

1.1.1.2 Infection, nodule organogenesis and type of nodules

The legume-rhizobial symbiosis results in *de novo* formation of specialized root organs, the nodules, that provide a suitable environment for N_2 fixation. In broad terms, the formation of symbiotic fixing nodules requires two developmental processes that are spatially and temporally coordinated: the nodule organogenesis, that proceeds in the plant cortex, and root hair bacterial infection (Roy et al., 2020).

Following the Nod Factors perception, root cortical cells reacquire properties typical of actively dividing meristem cells that will form the initial nodule primordium (Gibson et al., 2008) (Figure 2a). Simultaneously, as a result of successful bacterial colonization and initial nodule signaling, the high localized concentration of Nod Factors induces the curling of the root hair, which is followed by the invagination of cell wall trapping rhizobia and the formation of the infection thread (IT) (Gibson et al., 2008) (Figure 2a). From the trap site, the root hair begins and inverse tip growth, forming the IT tha is a tubular structure (Figure 2c). The IT continues to extend through the root hair until it joins the cell wall at the interface with the underlying cortical cell (Oldroyd et al., 2011). In this way, the IT carries rhizobia from the point of penetration toward the population of newly cortical meristematic cells (Gibson et al., 2008; Oldroyd et al., 2011). Ultimately, bacteria within the IT are completely internalized by root cells in an endocytosis-like process to form vesicles named symbiosomes, where they differentiate into a distinct cell type able to fix nitrogen, the bacteroids (Brewin, 2004) (Figure 2d).

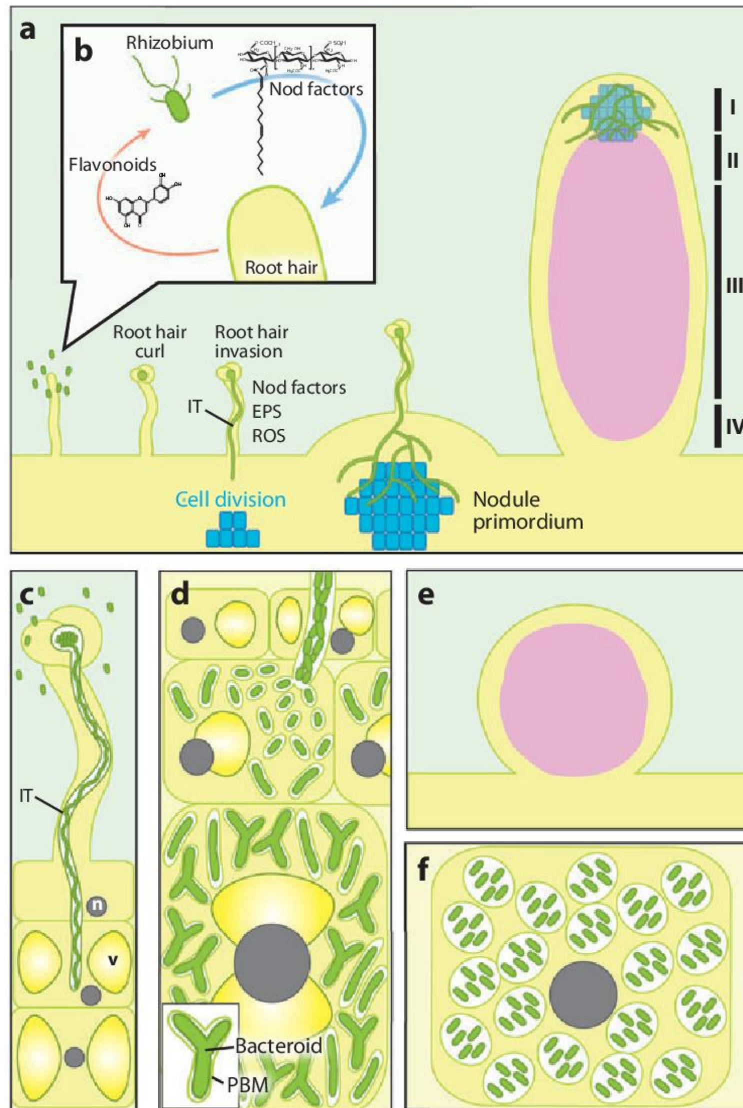


Figure 2. Schematic model of nodule development. (a-b) Host flavonoids exuded into the soil trigger bacterial Nod Factor production. Nod factor is perceived by host receptors and elicits various host responses, such as root hair curling and root hair invasion. Nod factors also induce mitotic cell division in the root cortex (represented in blue), leading to the formation of the initial nodule meristem. An indeterminate nodule originates from the root inner cortex and has a persistent meristem (Zone I). The nodule also contains an invasion zone (Zone II or Zone IZ) and a nitrogen-fixing zone (Zone III). In older nodules, a senescent zone (Zone IV) develops in which both plant and bacterial cells degenerate. (c) Bacteria enter the nodule through root hairs in a structure called the infection thread (IT) that elongates toward the nodule meristem; nucleus (n), vacuole (v). (d) At the tip of the growing IT, bacteria are endocytosed into the cytoplasm of meristematic cells. Each rhizobium is surrounded by a host-derived peri-bacteroid membrane (PBM) and proceeds to differentiate into the specialized symbiotic form called bacteroid. Bacteroids establish a chronic infection of the host cytoplasm and enzymatically reduce N_2 to provide a source of biologically usable nitrogen to the host (Zone III). (e) In contrast to an indeterminate nodule, a determinate nodule lacks a persistent meristem and all developmental stages proceed synchronously. (f) Infected cells of determinate nodules typically lack vacuoles (v). From Gibson et al. (2008), modified.

Generally, nodules fall into two morphological classes based on their pattern of meristem growth: determinate and indeterminate nodules, which occur in different species of legumes

(Gibson et al., 2008; Oldroyd et al. 2011). As occur in most temperate legume species (such as *M. sativa* and *P. sativum*), indeterminate nodules originate from inner cortical cells which form a persistent tip-growing meristem that continually produces meristematic plant cells (Gibson et al., 2008). Indeed, the plant, through the maintenance of the active apical meristem (which confer to the nodule the typical cylindrical shape) allows a prolonged bacterial invasion, following by the differentiation of both plant cells and endosymbiont (Gibson et al., 2008). As a result, a distal-to-proximal gradient of infected and uninfected cells along the structure is created (Gibson et al., 2008), and four distinct zones can be detected in a mature indeterminate nodule (Figure 2a). A fraction of undifferentiated rhizobial cells is maintained in the meristematic zone (Zone I), where new plant cells are produced, and in the invasion zone (Zone II) which contains infection threads full of rhizobia (Vasse et al., 1990; Poole et al., 2018). In the Interzone between zone II and III (Zone II/III or also identified as IZ) bacteria are released from infection threads and are internalized by plant cells. Instead, the nitrogen-fixing zone (Zone III) contains the mature N_2-fixing bacteroids, which degenerate together with the plant cells in the senescent zone (Zone IV) bacteroids senesce, respectively (Vasse et al., 1990; Poole et al., 2018).

In contrast, determinate nodules (as those of several tropical and subtropical legumes as *G. max*, *P. vulgaris*, and *L. japonicus*) (Gibson et al., 2008), derive from cell divisions of outer cortical cells and have a transient meristem. In the case of determinate nodules, the initial cell division activity required for nodule primordium formation ceases rapidly and therefore the determinate nodules contain no meristem. Accordingly, further nodule growth depends on an expansion in cell size rather than in cell number resulting in a spherical shape of the nodule (Puppo et al., 2005). Differentiation of infected cells occurs synchronously and the mature nodule contains symbiotic bacterial cells with a homogenous population of nitrogen-fixing bacteroids (Frassen et al., 1992). In general, the lifetime of the nodules is extremely variable, and it could depend on the legume species and on the environmental including stress conditions to which the plant is exposed (Van de Velde et al., 2006). Indeed, the nodule senescence is triggered prematurely by various types of stress (González et al., 1998; Matamoros et al., 1999). Differences in senescence are also present between the two types of nodules. Due to their typical spherical form, determine nodules are characterized by radial senescence of the tissue (Van de Velde et al., 2006). Contrarily, in the elongated indeterminate nodules, the senescence area (zone IV) gradually moves to the apical part initiating the degeneration of the nodule (Van de Velde et al., 2006).

Substantial differences occur in bacteroids maturation between the two types of nodules (indeterminate vs. determinate). Bacteroids in indeterminate nodules are functionally and morphologically distinct from free-living rhizobia undergoing a terminal differentiation process which is triggered by the plant host through the production of Cysteine-rich peptides (NCR) (Figure 4) (Mergaert et al., 2006). This profound differentiation process includes the genome amplification, generated by endo-reduplication cycles without cytokinesis, an increase in the permeability of the cell membrane, and results in an important cell swelling and an elongated shape (Vasse et al., 1990; Mergaert et al., 2006). These terminal differentiation program, however, precludes viability outside the host cell cytoplasm. Indeed, this differentiated bacteroids lose their ability to resume growth (Mergaert et al., 2006) and die with the nodule (Figure 3). However, undifferentiated rhizobia persist in the IT and can be released from the nodule tissue to the soil, giving rise to a new rhizobial soil population (Denison, 2000) (Figure 3). Differently, these bacteroids of determinate nodules are comparable to free-living bacteria in their genomic DNA content, cell size, and viability. Since these bacteroids are not terminally differentiated, they are released in the soil after nodule senescence and they can easily de-differentiate to the free-living form (Mergaert et al., 2006), enriching the rhizobia soil populations (Denison, 2000) (Figure 3).

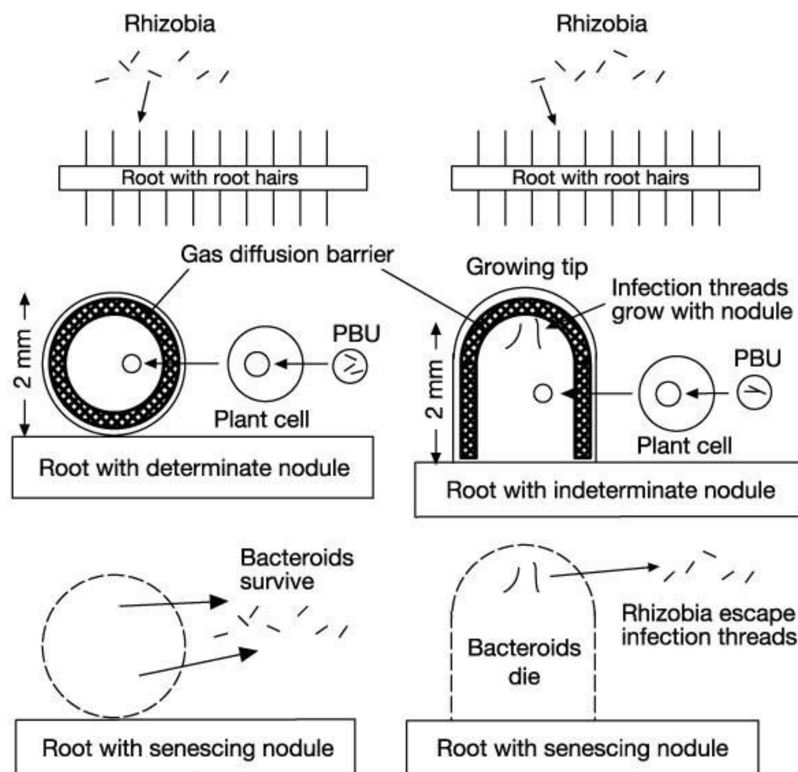
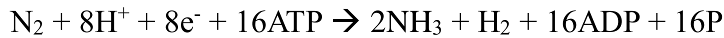


Figure 3. The Life cycle of rhizobia from soil to soil via determinate (left panel) or indeterminate (right panel) legume root nodules. Rhizobial differentiation into N_2-fixing bacteroids differs between the two nodule types. Bacteroids in determinate nodules resemble free-living rhizobia. Bacteroids in indeterminate nodules are usually larger and morphologically distinct from freeliving rhizobia; they are also individually surrounded by a peri-bacteroid membrane. When determinate nodules senesce, bacteroids survive and replenish soil populations. In indeterminate nodules, terminally differentiated bacteroids die, but undifferentiated rhizobia in the infection threads survive and reproduce in the soil. From Denison (2000), modified.

1.1.1.3 The nitrogen fixation process and nutrient exchanges

During differentiation from free living bacterium to bacteroid, a panoply of transcriptional changes occurs. Downregulation of many metabolic processes in conjunction with an increased expression of gene products involved in nitrogen fixation (Barnet and Fisher et al., 2006) happen and a global reviwing of carbon and amino acid metabolism take place (diCenzo et al., 2016; diCenzo et al., 2020). Indeed, inside the microaerobic environment of the nodules, the fully differentiated bacteroids express the oxygen-sensitive enzyme nitrogenase that catalyzes the conversion of atmospheric nitrogen to ammonia (Poole and Allaway, 2001) that will be further provided to the plant. The main factor that controlled N_2 fixation in bacteroids is low O_2 tension within nodules (Poole et al., 2018). The perception of low O_2 concentrations activates two oxygen-sensing bacterial regulatory circuits, which control the expression of specific genes

(*nif* and *fix*$^+$ genes) that codify for the nitrogenase complex and the micro-aerobic respiratory enzymes required to provide energy to nitrogenase (described in detail in Figure 4) (Fischer, 1994; Oldroyd et al., 2011). Indeed, the respiratory activity provides to nitrogenase 16 molecules of ATP and 8 electrons that are required to reduce 1 molecule of N_2 to 2 molecules of NH_3 (Poole and Allaway, 2001), according the following chemical reaction:



$$N_2 + 8H^+ + 8e^- + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P$$

Asides from the high amount of energy required, the BNF is strictly controlled since O_2 can inactivate the nitrogenase complex (Poole and Allaway, 2001). However, oxygen is also required for cellular respiration to supply the ATP necessary for the nitrogen fixation process (Poole and Allaway, 2001). To overcome this problem, the control of the microaerobic environment within nodules is mediated by the leghaemoglobin, an oxygen-binding protein that imparts a red color to functional root nodules and limits the concentration of free oxygen (Ott et al., 2005). This protein is composed of a heme group produced by rhizobial cells, and a globinic part provided by the plant (Poole and Allaway, 2001). Accordingly, due to its chemical structure, leghemoglobin can bind free-oxygen in the nodule facilitating the respiration of the bacteroids and simultaneously preventing the inactivation of nitrogenase (O'Brian et al., 1987).

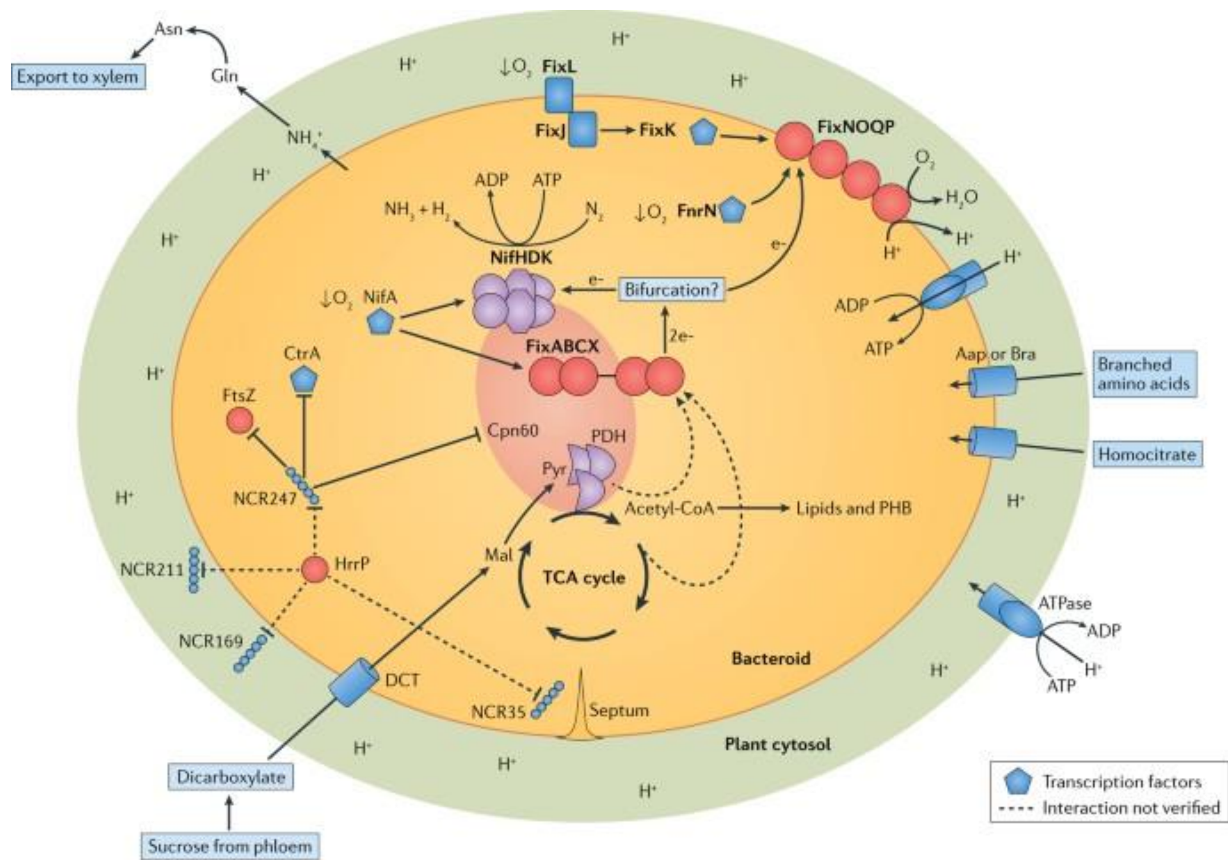


Figure 4. **O_2 is the main signal that regulates bacteroid development through two signalling circuits.** The first involves FixLJK in *S. meliloti* and probable transcriptional activator (FnrN) in *R. leguminosarum*, which induce the expression of *fix*NOQP. The second is Nif-specific regulatory protein (NifA), which induces *nif*HDK (encoding nitrogenase) and autoinduces *fix*ABCX–*nif*AB. FixABCX is an electron bifurcating complex, probably donating electrons to either ferredoxin or to flavodoxin and CoQ. Pyruvate dehydrogenase (PDH), the tricarboxylic acid (TCA) cycle and 2-oxoglutarate dehydrogenase might form a complex with FixABCX to achieve this, or, possibly, NAD(P)H provides electrons directly. FixNOQP (also known as CBB3) is a high-affinity terminal electron acceptor needed in O_2 -limited cells. These circuits are completely separate in *R. leguminosarum* but partly overlap in *S. meliloti* with weak induction of *fix*ABCX–*nif*AB by N_2 fixation regulation protein FixK. Low O_2 limits the TCA cycle, causing acetyl-CoA to be used to produce lipids and polyhydroxybutyrate (PHB). The peptidase HrrP is speculated to change host specificity by degrading nodule cysteine-rich (NCR) peptides. The HrrP–NCR peptide interactions in the figure are hypothetical and whether the shown peptides are important targets or in the correct compartment is unclear. The major chaperonin in bacteroids Cpn60 can interact with multiple proteins, including PDH. Aap, amino acid permease; Bra, branched amino acid permease; CtrA, cell cycle response regulator CtrA; DCT, dicarboxylate transporter; FixJ, transcriptional regulatory protein FixJ; FixK, N_2 fixation regulation protein FixK; FixL, sensor protein FixL; FtsZ, cell division protein FtsZ; Mal, malate; Pyr, pyruvate. From Poole et al. (2018), modified.

The BNF requires a large amount of energy which is provided by the host plant in the form of dicarboxylates, as L-malate (but also succinate and fumarate) (Figure 4 and Figure 5) (Poole et

al., 2018). The dicarboxylates are then transported across the symbiotic membrane inside the bacteroid, through the dicarboxylic acid transport (Dct) system, and metabolized via the tricarboxylic acid (TCA) cycle required to fuel nitrogen fixation (Lodwig and Poole, 2003) (Figure 4 and Figure 5). Ammonia (NH_3) is the primary stable product of the biological nitrogen fixation, it is crucial for the N assimilation for the plant, and it is transferred through bacterial cytoplasm to the plant (Lodwig and Poole, 2003) (Figure 5). Indeed, bacteroid avoid assimilating ammonia into an organic form, a process that is left to the plant which mainly assimilated the fixed nitrogen into glutamine and glutamate through the action of glutamine synthetase (GS) and glutamate synthase (GOGAT) (Udvardi and Poole, 2013) (Figure 5). However, according to some reports, L-alanine could also be an exported form of assimilated ammonia (Allaway et al., 2000). Therefore, fixed nitrogen is transferred from glutamine to asparagine or to purine derivates, as ureides, depending on the legume species (Vance, 2000). Temperate legumes (such as pea and alfalfa) which form indeterminate nodules, export mainly asparagine (Vance, 2000) (Figure 5). The switch on of ammonia synthesis by bacteroids (N_2 fixation) is accompanied by the switch off of ammonia assimilation into amino acids (Patriarca et al., 2002). Accordingly, bacteroids rely on the plant for the uptake of branched-chain amino acids, which are transported through symbiosome membrane by the Aap/Bra system (Figure 4 and Figure 5) (Prell et al., 2009; Prell et al., 2010). This phenomenon, ascertained in *R. leguminosarum* when in symbiosis with the plant, was named symbiotic auxotrophy and is then caused by the shutdown of amino acid synthesis in bacteroids (Prell et al., 2009; Prell et al., 2010).

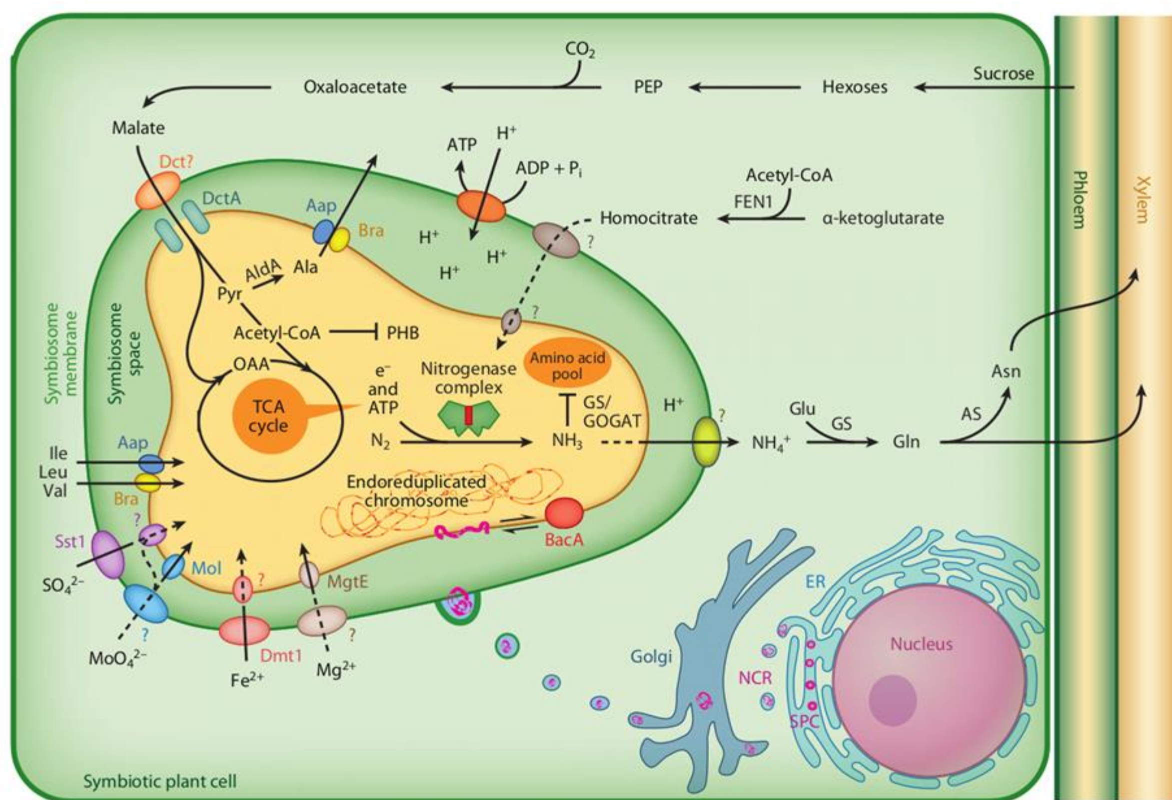


Figure 5. An idealized bacteroid in an indeterminate nodule. The bacteroid is enclosed in a plant-derived symbiosome membrane. Plants supply dicarboxylates (mainly malate, but also succinate and fumarate). Bacteroids shut down ammonia assimilation into amino acids and secrete ammonia to the plant. The transport systems identified which cause a Fix$^-$ phenotype are the putative Mg^{2+} transporter in *R. leguminosarum* MgtE (Karunakaran et al., 2009) and the sulfate transporter of *L. japonicus* SST1 (Krusell et al., 2005). Contrarily, the transport system divalent metal transporter DmtI present in *G. max*, is not essential for N_2 fixation (Kaiser et al., 2003). Developing or mature bacteroids become dependent on the plant for: i) provision of amino acids as seen by the requirement for Aap and Bra (Prell et al., 2009) ii) the FeMo Cofactor of dinitrogen reductase (enzyme II of nitrogenase) that requires homocitrate synthase of the plant (FEN1) (Hakoyama et al., 2009). Cysteine-rich peptides (NCR) are only made in indeterminate nodules of plants such as alfalfa, *M. truncatula* and pea but are lacking in *L. japonicus* and bean (Mergaert et al.,2006). From Oldroyd et al. (2011), modified.

1.1.2 Infecting legumes or forgoing symbiosis?

The intimate interaction between different organisms is classically named symbiosis (Leung & Paulin, 2008). The symbiotic relationships are commonly divided into three type of interactions, based on the benefits that these relationships may confer to the organisms involved: the mutualism, the parasitism, and the commensalism (Leung & Paulin, 2008). In the mutualistic

interaction, both the organisms involved receives benefit from the other, increasing their mutual fitness. conversely in parasitism only one partner benefits from the interaction while the other is affected negatively. In commensalism one species benefits while the other is unaffected. According to detailed researches, most of the global diversity is maintained by mutualisms– cooperative interactions among different species (Kiers et al., 2010; Bascompte and Jordano, 2007). Indeed, the mutualistic relationships ensure the survival and the reproduction of multitudes of organisms by providing essential ecosystem services such as pollination (Potts et al., 2010), seed dispersal (of several groups of plants, notably figs and yuccas) (Kiers et al., 2010), and critically participating to nutrient cycling, as for plant associated with fungal mutualists (Wilson et al., 2009; Kiers et al., 2010). Accordingly, the biological importance of mutualistic relationship is enormous.

The symbioses between legumes and the polyphyletic group of rhizobia, which includes members of alpha-proteobacteria and beta-proteobacteria (Vela´zquez et al., 2010), is a classic example of a mutualistic relationship, and is currently considered a model system for investigating the ecological and evolutionary aspects of mutualism (Kiers and Denison, 2008). Although rhizobia are textbook examples of cooperative symbionts, in nature they may cover the entire range of interaction, from mutualists (fixing symbionts) to parasites (non-fixing symbionts) varying dramatically in the N_2 sources provided to their hosts (Kiers and Denison, 2008). Moreover, infecting legumes and establishing the symbiosis (even providing little or no N_2) is not the only life strategy adopted by rhizobia. Free-living rhizobia strain exist, they lack of the genes for the infection of the plant, and are unable to nodulate legumes (non-symbiotic strategies). They are common and sometimes outnumber their related symbiotic strains (Denison and Kiers, 2004; Sullivan et al., 1996; Segovia et al., 1991; Laguerre et al., 1993).

Colonizing a nodule, and thus participating in symbiosis even with low or no contribution, can drastically increase the reproductive success of rhizobia. Indeed, the amount of reproductively viable rhizobia inside nodules greatly exceed their conspecifics in the soil. For example, a soybean nodule formed by a single strain may enclose up to 10^{10} descendants (West et al., 2002), a higher number compared to those of rhizobia contained in one gram of bulk soil (equal to 10^5 or fewer) (West et al., 2002). Although the chance for a single symbiotic rhizobium cell to successfully colonize a nodule of soybean is about one in a million (10^{-6}), an higher amount of viable descendants that can be produced inside the nodules (about 10^9), indicate that to infect plant is a positive life strategy for rhizobia (Denison and Kiers, 2004). Moreover, following the nodule senescence, these numerous reproductively viable rhizobia are released in the rhizosphere, increasing their densities in the soil and thus increasing the chance to infect a new

plant host (West et al., 2002). These observations are also valid for nodule with indeterminate growth, in which bacteroids lose the ability to reproduce. Nonetheless, in indeterminate nodules, large numbers of undifferentiated rhizobia persist in the infection thread, and are higher enough to escape into the soil and produce many descendants (Denison, 2000; Denison and Kiers, 2004).

Symbiotic rhizobia may thus have a greater fitness compared to a relative non-symbiotic that resides in the soil, where growth may be severely limited by nutrients availability. With fewer rhizobia that compete for the plant infection, symbiotic rhizobia would have greater average reproductive success than non-symbiotic strains. However, as the population of symbiotic rhizobia increases, the competition among relative rhizobia that can potentially infect the same plant host also increase, thus reducing the reproductive fitness of symbiotic rhizobia compared to those of non-symbiotic rhizobia (West et al., 2002). Though symbiotic rhizobia compete for host plants beside for the soil nutrients, they have not been displaced by the free-living rhizobia and, indeed, all the three life strategies, the symbiotic (mutualistic or parasitic) and non-symbiotic, evolutionarily coexist (Denison and Kiers, 2004). Others conditions that may affect the relative fitness of symbiotic *vs* non-symbiotic rhizobia (and among fixers and non-fixers strains) are reported in Table 2.

Table 2: Ovierview of the factors that affect the relative fitness of symbiotic *vs* nonsymbiotic. The fitness benefits to a symbiotic rhizobium from attempting to nodulate a legume is the product of i) the chance of success, ii) the average number of viable rhizobia released from a senescing nodule, and iii) the probability that a given rhizobium will survive long enough after nodule senescence to become established in the soil. The table summarizes factors affecting these three parameters (first line) and how they, in turn, affect the relative fitness of symbiotic *vs* non-symbiotic rhizobia and (among symbiotic rhizobia) *vs* less-mutualistic strains. From Denison and Kiers (2004), modified.

	Chance of founding a nodule > if:	**Rhizobia released per nodule > if:**	**Survival after leaving nodule > if:**
Symbiotic rhizobia have greater fitness than those that never infect legumes if:	Symbiotic rhizobia are rare, there are more host plants, or less soil nitrogen	Fewer symbiotic rhizobia and less soil N may favor nodule growth	Rhizobia hoard PHB in nodule – but nodule may attract predators
Among symbiotic rhizobia, those that fix N_2 have greater fitness if:	Some non-fixing rhizobia may be recognized and excluded	Host sanctions favor N_2-fixers, but maybe less if mixed nodules common	Lingering effects of sanctions exceed benefit from maybe saving more PHB

1.1.3 Cooperation and conflicts of interest

Classically, the mutualistic symbiosis established among legumes and rhizobia should provide notable benefits for both partners. However, this interaction also implies costly biological process for both the protagonists, nitrogen fixation for rhizobia and nodule organogenesis for plant (Denison and Kiers, 2011). Each partner has also to allocate the achieved resources between its own growth and reproduction and the service it provides to the other (De Mita, 2012; Denison 2000). The photosynthetically fixed carbon compounds, that they could employ to uphold the rhizobia reproduction, are instead metabolized to support the expensive process of the N_2 fixation. Accordingly, rhizobia face a “resource allocation trade-off” between the employment of plant-provided sugars to boost the N_2 fixation, and hoarding these resources for their future survival in the soil. Similarly, plants may re-modulate their C allocation to rhizobia in different growth stages, i.e. during seeds development, or with seasonal changes, as perennials plants during winter (Denison, 2000). Therefore, partners may experience fitness conflicts of interest that may lead to the mutualism breakdown and drive an antagonistic behaviour (Sachs and Simms, 2006; Friesen, 2012). For example, ineffective rhizobia induce

the nodule formation but fail in N-fixation establishing a parasitic interaction with the plant. Indeed, some aspects of the legume-rhizobia symbioses may foresee the possibility of conflicts of interest, in which:

i) Several rhizobia strains displayed a variation in efficiency in nitrogen-fixation among different host genotypes (Thrall et al., 2007; Heath and Tiffin, 2007; Heath and Tiffin, 2009; Sachs et al., 2010; Regus et al., 2017). Moreover, even if isolated from a single location and thus adapted to the same soils condition, the net benefits provided to the host by different strains can vary up to tenfold (Burdon et al., 1999).

ii) Rhizobia spread among hosts through the soil. Differently from vertically-transmitted mutualisms, in the horizontal transmission the offspring of both partners are not correlated, implicating that the rhizobial fitness may not be fully aligned with the fitness of the plant (Heath and Tiffin, 2009; Westhoek et al., 2017).

iii) The controversial question of the existence (and efficiency) of the partner choice mechanism. Few studies inferred that legumes were able to exclude parasitic rhizobia. The 'partner choice' was thus considered a robust mechanism exerted by plants against the entry of ineffective strains in nodules before the establishment of symbiosis (Heath and Tiffin, 2009; Gubry-Rangin et al., 2010; Regus et al., 2014). Nevertheless, these studies were mainly conducted with effective and ineffective mixtures of non-isogenic strains, without considering the different competitiveness in plant roots colonization. A recent study has demonstrated that the plant is unable to discriminate between effective and ineffective strains before the nodule formation (Westhoek et al., 2017).

iv) Ineffective strains are known to be common in field conditions (Denison et al., 2004). Indeed, these highly competitive strains are generally able to replace effective strains employed as bio-inoculants, thus causing a decrease in crops yield (Kosslak et al., 1983; Triplett and Sadowsky, 1992; Kiers et al. 2007).

1.1.4 Features of ineffective strains

As long as to a greater N_2 fixation correspond a higher amount of carbon sources, rewarded by the plant that can be further exploited by rhizobia, natural selection should favor strains whose bacteroids fix the most N_2 and thus promote the symbiotic cooperation (Denison, 2000). However ineffective rhizobia are widespread and represent a common problem for agricultural

legumes (Kiers et al. 2007, Westhoek et al., 2017). A common factor among the strategies adopted by ineffective strains, to support their own subsequent growth in the soil, is the ability of hoarding plant resources (Ratcliff et al., 2008). Carbon supplied by the plant may be stored in bacteroids as poly-β-hydroxybutyrate (PHB). Indeed, large amounts of PHB are usually accumulated by bacteroids in determinate nodules acting as a store of carbon to use after the nodule senescence (Lodwig et al., 2005; Terpolilli et al., 2016). The PHB can also be detected in undifferentiated rhizobia (which have not lost the ability to dedifferentiate into free-living bacteria once they are released in soil) inside the infection thread of indeterminate nodules (Bergensen et al., 1991; Lodwig and Poole, 2003), where it can be degraded to fuel the bacteroid differentiation process (Lodwig et al., 2005, Jiabao Xu et al., 2017). However, as these undifferentiated cells of indetermine nodules are released into the bulk soil, this accumulated high-energy molecules could also support rhizobial survival and provide a competitive advantage over the other bacteria (Trainer and Charles, 2006). To ensure their reproductive success after nodule senescence, non-fixers strains may be particularly able to divert carbon sources into PHB (Denison & Kiers, 2004). Indeed, mutants unable to synthesize PHB showed higher N_2 fixation rates than their parental strains (Cevallos et al., 1996). Similarly, an inverse proportionality between the N_2 fixation and PHB production was also verified in non-fixing mutants (Hahn and Studer, 1986). These evidences indicate a clear trade-off between the two processes (N_2-fixation and PHB accumulation) (Friesen, 2012). Accordingly, the PHB accumulation works as a sink of a source of energy for free-living rhizobia that would otherwise be employed in symbiotic interaction, specifically for nitrogen fixation (Muller and Denison, 2018).

Bacteroids infecting indeterminate nodules, as *S. meliloti*, are known to divert the obtained resources in the synthesis rhizopines, which are usually provided to their related undifferentiated rhizobia cells as a source of carbon (Murphy et al. 1995). Besides the synthesis of PHB, the production of rhizopines may also be considered a cheating strategy. Indeed, in nodules singly infected by ineffective rhizobia, rhizopines production may be at the expense of N_2 fixation, thus compromising this process (Oono et al., 2009). Though the first hypothesized role for rhizopines was in the kin selection of related rhizobia in the rhizosphere (Olivieri and Frank, 1994; Simms and Bever, 1998), the production of these molecules is rather involved in a within nodule kin selection. This type of kin selection is likely the most powerful form of selection that consolidates the cooperation between partners (Oono et al., 2009), and ensure the increase of the fitness and the persistence of ineffective rhizobia. However, even without

resource diversion via rhizopines or PHB, still available sources may be redirected to the undifferentiated rhizobia simply by diffusion (Oono et al., 2009).

1.1.5 The “Tragedy of the Commons” and a new concept of cheating

In natural settings, the issue of conflicts of interest between partners is further complicated by the evidences of multiple infection of the host nodules by different rhizobial strains (Dennison and Kiers, 2004; Westhoek et al. 2017). In mixed infections, cheater strains can gain advantage by hoarding plant resources and storing energy for survival in the soil, while leaving the costly process of nitrogen fixation to others symbionts (Heath and Tiffin, 2009; Kiers and Denison, 2008). Basically, less effective strains could freeload on the sources generated by the mutualism with more effective strains (Denison, 2000; Friesen and Heath, 2013; Kiers et al., 2013). Therefore, simultaneous infection by multiple strains can introduces a symbiont-symbiont conflict, in addition to the current host-symbiont conflict, and may alter the evolutionary dynamics of the interaction (Friesen and Mathias, 2010).

The “tragedy of the commons” scenario theorized by (Hardin, 1968) seems to be particularly well-suited for legumes-rhizobia dynamics, in which “the public goods” are represented by the host sources that can be subject to exploitation of potential partners (Friesen and Mathias, 2010). In this sense, when several rhizobial strains infect the same plant, the “tragedy” is that cooperative partners (N-fixers) that supply their hosts with nitrogen, indirectly support other rhizobia strains (including non-fixers ones) competing for host resources and their future fitness (Denison et al., 2003a; Kiers and Denison, 2008) (Figure 6).

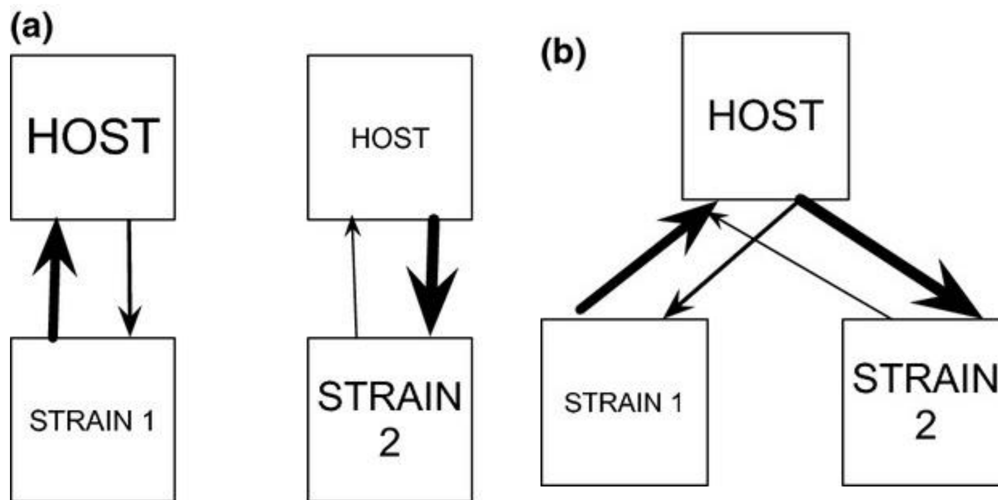


Figure 6. Direct cheaters and cheaters of cooperative strains. (a) Direct cheaters (strain 2) that increase their fitness benefit at the expense of host fitness would be detectable by single-strain inoculations when compared with a cooperative strain (strain 1). (b) Cheaters that exploit the benefits generated by a cooperative strain would only be detectable in multi-strain inoculations. From Friesen and Heath (2013), modified.

Hence, thinking at the concept of cheating as referred to the interaction symbiont-symbiont infecting the host instead of the sole interaction symbiont-plant, the definition of "cheaters" as "individuals who do not cooperate (or cooperate less than fair share) but potentially able to gain the benefit of others cooperative partners" (West et al., 2007) seems to be more appropriate. A *B. japonicum* strain was identified as cheater able to proliferates without providing no growth benefits to host plants, reaching a higher density than more effective strains in nodules (Sachs et al. 2010a). Moreover, when competing against beneficial strains, this strain forms few and small nodules (Sachs et al. 2010b), typical of the infections with non-fixing rhizobia (Singleton and Stockinger, 1983; Kiers et al., 2003, Simms et al., 2006; Regus et al., 2014, 2015; Westhoek et al., 2017) (Figure 7). The cheater *B. elkanii* strain USDA61, interfering with plant ethylene signaling, can accumulate greater amounts of PHB molecules and, thus, increase their own fitness (Ratcliff and Denison 2009; Kiers et al. 2013).

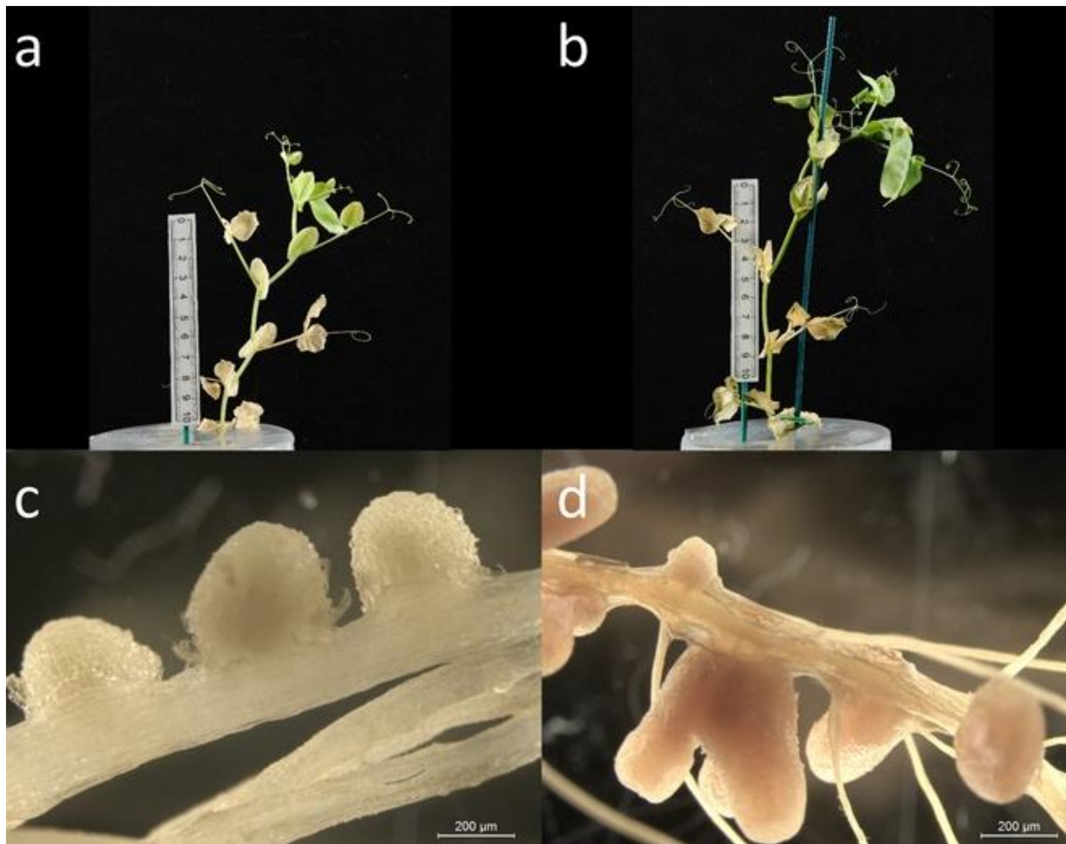


Figure 7. Comparison of pea plants inoculated with non-fixing *R. leguminosarum* bv. *viciae* strain RU3940 and the isogenic fixing *R. leguminosarum* bv. *viciae* strain Rlv3841. Nodules and pea plants inoculated with the non-fixing *R. leguminosarum* bv. *viciae* strain RU3940, (a, and c) and the otherwise isogenic fixing *R. leguminosarum* bv. *viciae* strain Rlv3841 (b and d). After five weeks of growth, plants inoculated with the non-fixing strain (a) had reduced biomass and produced fewer peas than those inoculated with the fixing strain (b). The non-fixing strain formed white (c), rather than pink (d), nodules, indicating a lack of leghaemoglobin. From Westhoek et al. (2017), modified.

1.1.6 Host sanctions and other stabilizing evolutionary mechanisms

Evolutionary biologists have questioned which are the main evolutionary forces that drive a mutualistic interaction and its stability. In particular, the issue of cheating, which is nowadays regarded as a focal concept in the study of mutualism (Jones, 2015), poses a fundamental problem to the evolutionary persistence of cooperation and captured the attention of biologists. As mentioned above, legumes are typically infected by several strains. In the absence of a mechanism to prevent parasitism, in the tragedy of the common cheater strains might be expected to displace cooperative strains. Consequently, in the rhizobia-legume symbiosis a key question is: why do rhizobia expend energy to provide benefits to their host plant (fixing N_2)

instead of using it for their own reproduction and survival? (Marynard Smith 1989; West et al., 2002). In other words, why do rhizobia fix nitrogen?

Two potential mechanisms were proposed to stabilize the cooperative behavior of the bacterial symbionts at the infection and/or post-infection levels: “partner choice” and “host sanctions”. Partners choice is usually defined as “a mechanism that allows the detection of suitable partners before the establishment of the infection” and based on signal recognition (Westhoek et al., 2017; Daubech et al., 2017). As mentioned in section 1.4.2, partner choice is nowadays considered unsuitable in preventing the infection by ineffective rhizobia, since the inability of the plant to assess the quality of the partners in advance (especially if fixers and non-fixers strain are closely related). After all, rhizobia do not have any reason for warning the host about their inefficiency (Masson-Boivin and Sachs, 2018). However, a sort of partner choice at pre-infection stages in *M. truncatula – S. melioti* system was hypothesized (Gubry-Rangin et al., 2010). In contrast to this pre-infection stage mechanism, post-infection sanctions rely on the ability to discriminate between low- and high-quality partners after the nodule organogenesis, and to punish or reward them accordingly (Frederickson, 2013). In particular, sanctions were defined as “some action by the host plant that reduces the relative fitness of less beneficial rhizobia based on their low rate or low efficiency of N_2 fixation” (Oono et al., 2009). In the model proposed by West and collaborators (2002), appreciable levels of N_2 fixation are only reached in presence of plant sanctions that target low-quality partners at the whole-nodule level, thus supporting “the sanction hypothesis” previously suggested by Denison (2000). Accordingly, the N_2 fixation process becomes an evolutionarily stable strategy for rhizobia, and the sanction a potential stabilization mechanism for cooperation between partners that prevent losing resources in sustaining cheater strains (West et al., 2002; Kiers and Denison, 2008). In this perspective, the reciprocal exploitation conceived in mutualism turns into a “balanced exploitation”, in which the dominant partner enforces “good behavior” on the other (West et al., 2002). Beyond the general reducing (or re-allocation) resource losses in nodules infected by less-mutualistic strains, other metabolic adjustments were proposed as sanctions. The upkeep of the micro-aerobic environment inside nodules is critical for the N_2 fixation, to avoid oxygen inactivation of nitrogenase (Udvardi and Poole, 2013). However, rhizobia are obbligate aerobes that require oxygen for respiration and energy metabolism. Therefore, facing this dilemma under nitrogen-fixing conditions, concerted efforts of both plants and rhizobia to produce a microaerobic environment around nitrogen-fixing rhizobia in nodules are exerted (Udvardi and Poole, 2013). Besides leghemoglobin that bind oxygen, a barrier to gaseous diffusion in the nodules’s outer cell layers limits the rate of oxygen influx to the central infected

tissue. Using argon gas to prevent N_2 fixation, it was demonstrated empirically that a decrease in O_2 influx to nodules may occur severely reduces the growth rate of rhizobia (Kiers et al., 2013, 2006; Oono et al., 2011). Indeed, a decrease in host-mediated O_2 supply to the nodule interior coincided with a 50% reduction in rhizobium reproduction in the non-fixing nodule (Kiers et al., 2003). A decrease in nodule O_2 permeability was hypothesized to limit rhizobium reproduction (Kiers et al. 2003), either directly or indirectly (e.g., via an O_2 effect on C supply to rhizobia from the plant cells in which they are found). Moreover, host plants can also induce an accelerated schedule of nodules senescence for those that are infected by rhizobia that fix negligible nitrogen (Sachs et al., 2010b; Oono et al., 2011; Regus et al, 2015). Indeed, if nitrogen fixation is insufficient, the plant may interrupt the symbiosome development through the neutralization of the peri-bacteroid space (Pierre et al., 2013; Sachs et al., 2018).

In legume-rhizobia symbioses, partners can also invest their resources at intermediate levels between cheating and cooperating. Indeed, some rhizobia may modulate their fixation rate to prevent possible sanctions (Denison and Kiers, 2004). Plant control could also be sophisticated: it is conceivable that intermediate sanctions exist, as was shown occur in soybean (Kiers et al., 2006), allowing the evolution of intermediate levels of cheating (Kiers and Denison, 2008). Moreover, was demonstrated that legumes may regulate sanctions against specific nodules based on the total amount of supplied nitrogen (West et al., 2002). However, most of the evidence of host sanctions was observed in determinate nodules (Gubry-Rangin et al., 2010), in which bacteroids lose the ability to dedifferentiate. Furthermore, in some interaction (as in *Glycine max* - *Bradyrhizobium japonicum*, and *P. vulgaris – R. etli*) host sanctions were not detected at all (Marco et al., 2009; Ling et al., 2013). Another post-infection mechanism called partner-fidelity feedback (PFF) and exerted by the plant host is to reward the symbionts based on their mutualistic performances. This positive mechanism result in differential nutrient influx toward nodules according to their ammonium output. During long-lasting interactions, the PFF may also strengthen the positive selection of effective symbionts in spatially structured environments, hereafter of signal recognition or performance evaluation (Sachs et al., 2004). Given the heterogeneity of mechanisms among legumes that establish the symbiosis, it is thus reasonable to infer that, that different plants may rely on different control mechanisms (Masson-Boivin et al., 2009).

1.1.7 Why do cheater rhizobia still exist?

The possibility that plants can sanction unproductive nodules, or to preferentially allocated sources to those containing fixing strains (Oono et al., 2009), is commonly recognized. If every nodule was founded by a single strain, sanctions imposed at the level of individual nodules (termed "whole-nodule" sanctions) would reduce the relative fitness of less-mutualistic strains. However, in nature, the cheating polymorphism still persists and, above all, evidence of mixed nodules in field-grown host plant (Moawad and Schmidt, 1987; Simms et al., 2006; Van Berkum et al., 2012), and in multiple legume species in laboratory conditions (Brockwell et al., 1977; Amarger, 1981; Gage, 2002; Heath and Tiffin, 2009 Sachs et al., 2010b; Ling et al., 2013; Checcucci et al., 2016, Daubech et al., 2017; Mendoza-Suárez et al., 2019) were observed. Hence, the prospect of mixed nodules complicates the scenario described so far.

The chance for cheaters of sharing the same nodule with effective strains (mixed nodules) may represent an attractive opportunity to escape plant sanctioning (Kiers and Danison, 2008), then for proliferating without detection (Friesen and Mathias, 2010, Checcucci et al., 2016). Indeed, according to a mathematical model, in mixed nodules a decrease of the efficacy of sanctions is predicted, and the exploitation of mutualistic strains by cheaters increases with the co-infection rates (Friesen and Mathias, 2010). *R. etli* mutants in *nodB* gene, involved in the deacetylation of Nod factors in rhizobia, showed the ability to invade host nodules if co-inoculated with their wild type, foreseeing the possibility of a ''hitch-hiking'' strategy adopted by cheaters during the colonization process (Ling et al., 2013). In *M. sativa – S. meliloti* interaction, the widespread evidence of mixed nodules, founding by both mutualistic and cheater strains, seems not particularly affect the plant growth (Checcucci et al., 2016). Similarly, the co-infection of *P. vulgaris* with *R. etli nifA* mutants and their wild type does not alter N_2 fixation rates (Ling et al., 2013). Therefore, in mixed nodules, in attempts to avoid "whole-nodule sanctions", mutualistic strains may increase their N_2 fixation performances to balance the nitrogen deficit caused by the presence of the cheaters.

Mixed nodules may also exhibit sectored patterns consistent with the co-infection of more strains (Figure 8). Indeed, a nodule spatial separation of bacteroids was observed in mixed nodules of *Lotus strigosus* and *Lotus japonicus* both infected with a mixture of effective and ineffective strains (belonging to *Bradyrhizobium sp*. and *M. meliloti*, respectively) (Regus et al., 2017). Likewise, fluorescent sectoring was detected in alfalfa nodules founded by GFP and RFP-expressing *S. meliloti* strains (Gage et al., 2002), and in *M. pudica* nodules. The natural sectoring of the symbionts population in mixed nodules may allow mechanisms of cell-

autonomous sanctions within determinate co-infected nodules, hitting only less-effective rhizobial genotypes without disrupting the integrity of the entire nodule (Regus et al., 2017). Similarly, *Cupriavidus taiwanensis* uncooperative strains do not persist within cells of indeterminate nodules of *M. pudica*, even if shared with fixing symbionts, probably as a result of a cell-autonomous senescence program (Daubech et al., 2017). However**,** these "targeted-sanctions" may be exclusively direct toward non-effective bacteroids of the nitrogen fixation zone, allowing released undifferentiated rhizobia to persist in soil and, more importantly, not extinct within (mixed) nodules (Daubech et al., 2017). Hence, this "loose selection process" coupled with partial overlap of the nodulation and nitrogen fixation abilities, can retain the genetic diversity of rhizobial communities in the soil and shape the ecology and evolution of rhizobia (Moisson-Boivon and Sachs, 2018).

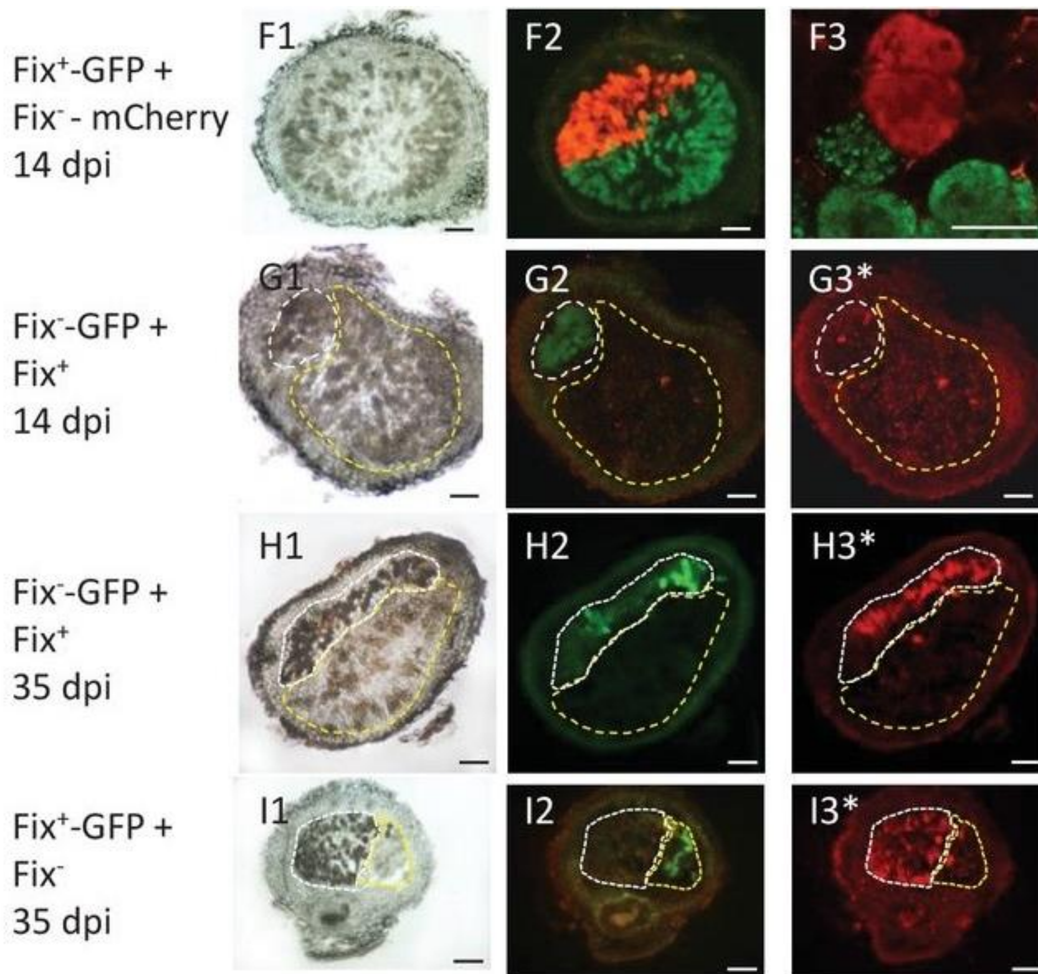


Figure 8. Spatial separation of N-fixing (Fix$^+$) and N-non-fixing (Fix$^-$) bacteroids in mixed nodules of *Mimosa pudica*. *M. pudica* were co-inoculated with Fix$^+$ and Fix$^-$ *C. taiwanensis* at a 1/1 ratio and sections of nodules collected at 14 dpi (FG) or 35 dpi (HI) were observed under bright field (panels 1) or fluorescent microscopy (panels 2 and 3), and after PI staining (panels with an *). Panels with the same letters represent the same nodule section. (F3), magnification of (F2) visualized by confocal microscopy. Sections: (F) nodule co-infected with a GFP-labeled Fix$^+$ and a mCherry- labeled Fix$^-$ strain. (G) and (H), nodules co-infected with a GFP-labeled Fix$^-$ and an unlabeled Fix$^+$ strain. (I), nodules co-infected with a GFP-labeled Fix$^+$ and an unlabeled Fix$^-$ strain. The white and yellow dotted lines in (GHI) delimit the areas occupied by the Fix$^-$ and Fix$^+$ strains in a co-infected nodule, respectively. Note that the Fix$^-$ bacteroids (G3) are red-labeled by PI staining at 14 dpi, and Fix$^-$ are mostly PI-labeled (dead) at 35 dpi (H3I3). Note that nodule cells filled with Fix$^-$ are browner than nodule cells filled with Fix$^+$ (G1H1I1). Scale bars correspond to 100 mm except for F3 (30mm). From Daubech et al. (2017), modified.

1.1.8 Single-strain inoculation *versus* co-inoculation

Since legume-rhizobial symbiosis is widespread, it is evident that symbiosis confers a fitness advantage for both partners (Remigi et al., 2014; 2016). Generally, N_2 fixation mutualistic traits are favored in rhizobia during the evolution, as also confirmed by experiments in laboratory conditions (Kiers et al., 2003; Simms et al., 2006; Sachs et al., 2010b; Regus et al., 2014; Regus et al., 2015; Masson-Boivin and Sachs, 2018). Nevertheless, uncooperative rhizobia (non-fixing strains) persist in a variety of settings (Masson-Boivin and Sachs, 2018). More importantly, the relevance of cheaters on plant fitness (that is, determine if cheaters can really affect plant growth) is still debating and has to be clarified (Checcucci et al., 2016). Therefore, the question of whether legume fitness and rhizobia fitness are truly aligned is yet open. The scenario presented so far highlighted several processes (resource hoarding by cheater strains, the existence of host sanctions and mixed nodules) underlying legume-rhizobium fitness alignment or conflict (Figure 9). A non-alignment of fitness of partners, thus the opportunity of antagonistic co-evolution, mainly depend on the infection rates, determined by the competition for nodulation, and resource exchanges (Friesen, 2012). In particular, nodulation competitiveness directly influences a strain's relative fitness in a population (Friesen 2012). Indeed, successful colonization due to good competition abilities increases the reproductive success of rhizobia.

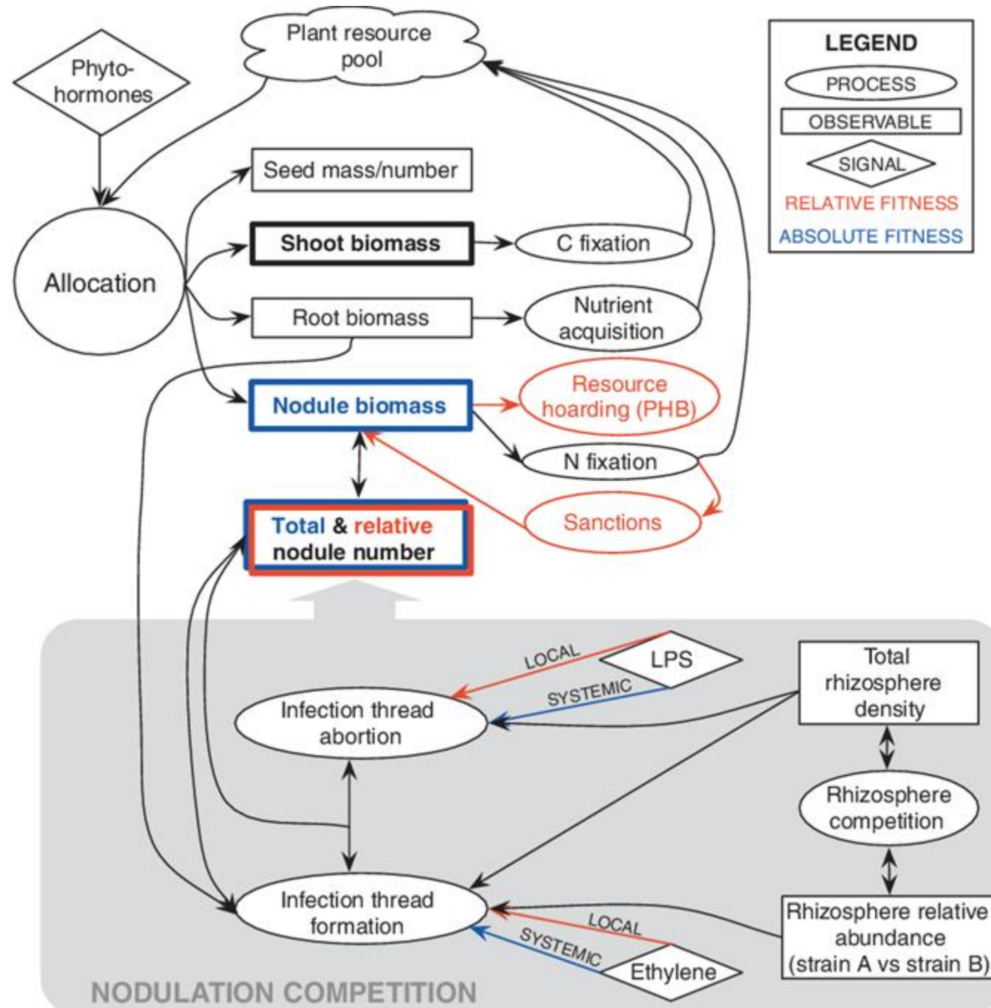


Figure 9. Simplified diagram of processes underlying legume and rhizobia fitness alignment and conflict. Processes (ovals) and relationships (arrows) are separated into those that represent direct feedback to the relative fitness of a strain (red) vs feedbacks that influence absolute fitness of all strains infecting a plant (blue; this could include bacterial cell number and cell quality as long as benefits are distributed equally across strains). Rectangles are observable quantities; bold rectangles represent the most commonly measured aspects of the interaction: aboveground biomass, nodule biomass, nodule number and relative nodule number. Signals modulate specific processes, such as phytohormones altering plant allocation patterns. Signals that influence infection thread formation (reviewed in Gage, 2004) include rhizobium-produced nod factor (reviewed in Long, 1996), succinoglycan (Jones et al., 2008) and plant-produced molecules that regulate root hair infection, including CLE peptides, ethylene, jasmonic acid and abscisic acid (Penmetsa & Cook, 1997; Ding et al., 2008; Ferguson et al., 2010; Batut et al., 2011). Many of these plant-derived molecules are known to act both locally and systemically (e.g. Mortier et al., 2010). Several plant genes required for infection thread development have been identified (Popp & Ott, 2011), and both nod factor (Walker & Downie, 2000) and bacterial surface molecules (lipopolysaccharide (LPS) and succinoglycan) can regulate infection thread abortion (Dazzo et al., 1991; Cheng & Walker, 1998). PHB, polyhydroxybutyrate. From Friesen (2012), modified.

However, the nodulation competitiveness is not frequently considered in the evaluation of the reproductive rhizobia fitness, which is often assessed by using single-strain inoculation. Indeed, by using single-strain inoculation, the estimated (absolute) fitness does not correspond to the real fitness (relative fitness) of rhizobia in competitive conditions (Oono et al., 2009) and it might be overestimated as its competitive abilities. This problem is highlighted the case of a Californian strain of *Bradyrhizobium* that showed high absolute fitness when inoculated alone, forming many nodules with high bacterial cell numbers, but extremely low relative fitness in competitive conditions (Sachs et al., 2010a; 2010b).

To address the issue of the alignment of fitness, a recent meta-analysis was conducted by Friesen (2012) in which positive correlation between a strain’s competitive success in nodulation and its impact on plant growth was founded. In this work, the plant-benefit data were all collected by singly-inoculated plants as well, and no measures of direct fitness of both partners (as number of seeds and number of rhizobial cells released) were reported. However, higher competition for nodulation ability doesn’t imply a greater N-fixation efficiency, since they are independent traits (Kiers et al., 2013). Therefore, also in this case, single strain inoculation can enhance the competition abilities, especially for fast-nodulating strains, but it might minimize the N_2 fixation efficiency of lower-nodulating strains (Kiers et al., 2013).

Single-strain inoculation represents one of the greater constraints to address the issue of fitness alignment among partners, as spurious correlations between plant fitness and rhizobial fitness may arise (Kiers et al., 2013). This is particularly evident when cheater and a cooperative strain are compared employing both single-strain inoculation and co-inoculation (Figure 10). Indeed, co-inoculation with the cheater and cooperative strains would presumably lead to intermediate plant growth, or poorly reduced, compared to those with the two strains separately (Kiers and Denison, 2008; Kiers et al., 2013). Hence, intermediate levels of effectiveness in promoting plant growth cannot be detect by using singly inoculation (Figure 10). Moreover, in co-inoculation, the rhizobia fitness differs from the single-strain inoculum showing how the fitness of cheater strains may be higher than the cooperative ones (Figure 10). This picture partially explains why highly beneficial inoculum strains tend not to persist in field over years (Kiers and Denison, 2008), demonstrating that the selection of possible rhizobial inoculants, based exclusively on ascertained performance with single-strain inoculation, could not be a successful strategy.

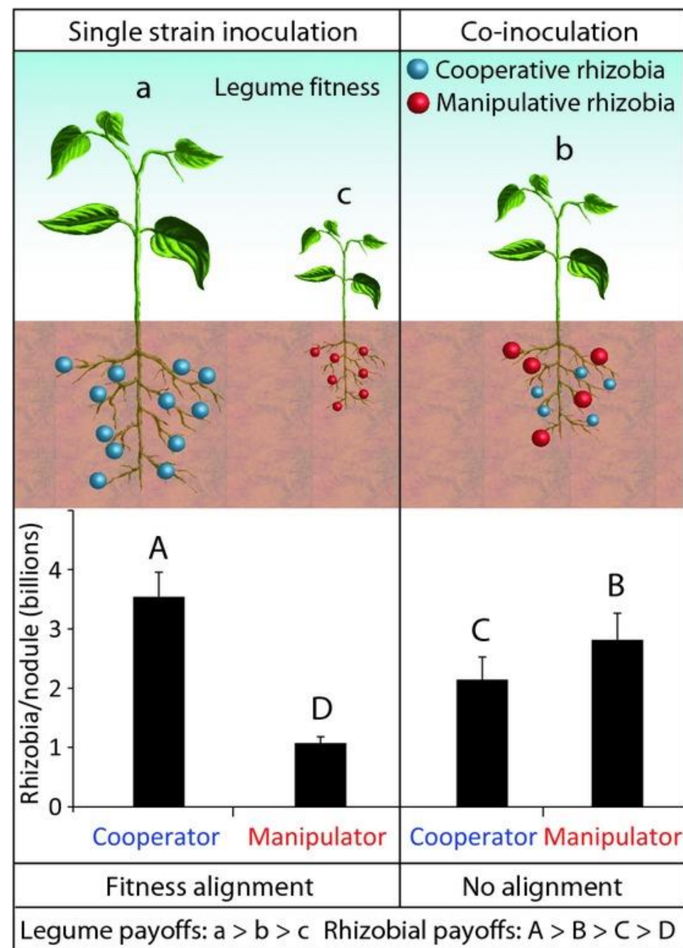


Figure 10. Legume fitness and rhizobia fitness in single-strain inoculations and co-inoculations. Under single-strain inoculation, more cooperative rhizobia directly benefit from greater plant growth, resulting in higher host and symbiont fitness (left). Any rhizobia that provide less host benefit and yet evade sanctions do poorly when alone, but exploit cooperative rhizobia when they co-infect a host (right). Data from Ratcliff & Denison (2009) are here elaborated. The number of rhizobia per nodule, corrected for potential reproduction from stored PHB, of the cheater *Bradyrhizobium elkanii* Rtx (+) (that greatly accumulate poly-hydroxybutyrate (PHB) and of the isogenic Rtx (−) mutant, either singly infected or co-infected hosts, are shown. The potential reproduction from stored PHB was calculated as described in Ratcliff et al., (2012). From Kiers et al. (2013), modified.

1.2 The ecology of competition

Considerable progress in studying the conceptual issues and mechanisms surrounding the ecology and evolution of mutualism were witnessed last years (Jones et al., 2012), some of them deeply discussed in the previous paragraph (i.e. how is mutualism maintained and how partners co-evolve). Mutualism can be described as consumer–resource interactions (Holland et al., 2005), based on the exchange of resources among partners acting as service or reward

(Bronstein, 2009). In this perspective, referring to legume-rhizobia symbiosis, the competition for the plants produced resources among putative partners takes place, and gaining access to these benefits can thus confer a competitive advantage to the 'winners' ecologically and evolutionary speaking. Different competitive scenarios can occur in the race for the shared partner (or shared sources), as illustrated in detail in Jones et al. (2012), which may not depend directly upon the mutualistic-quality of competitors. Specifically, in the case of competition between mutualists and cheaters strains, plant control mechanisms (as sanctions) should influence the competitive outcomes in favor of the latter. Nevertheless, the existence of mixed nodules, as well known, may challenge this assumption and the more competitive abilities for an initial nodule occupancy of the cheaters may play a central role in competitive dynamics and in competitive outcomes. However, the strength and consistency of such competitive abilities of cheaters over the mutualists and vice versa, have been yet examined (Jones et al., 2012). Besides the narrowly ecological definition of competition meaning as competition for plant sources, that may chiefly take place within nodules, what exactly is the definition of competition? The term "competition" mainly referred to the dynamic that is established when several compatible rhizobial strains (in the rhizosphere or near the root system) "compete" for the nodules occupation (Amarger 1981; Thies et al. 1991b). In this regard, in mixed infections the strain that showed the higher nodule occupancy is likely labeled as the more competitive (Yates et al., 2011).

1.2.1 Features that affect nodulation competitiveness in rhizobia

The competition dynamics may occur at a different level (or steps) reflecting the different phases that rhizobia have to overcome for occupying a nodule: from the rhizosphere colonization, the adhesion to roots surface, the infection process and the nodule development, to nodule colonization and differentiation into N_2-fixing bacteroids (Oldroyd et al., 2011; Wheatley and Poole, 2018; Iturralde et al., 2019). Over time, several phenotypic traits involved in different steps (the rhizosphere colonization, the rhizobial adhesion to root surfaces and the infection process) were recognized as significant for rhizobia competitiveness.

Motility skills are necessary during the rhizosphere colonization. A decrease of roots nodulation was observed in motility defective mutants of *S. meliloti* and *R. leguminosarum* (Caetano-Anollés et al., 1988b; Mellor et al., 1987; Parco et al., 1994). However, also external factors may influence the rhizobia motility towards root surfaces, which is limited in field (López-

García et al., 2002), thereby positively affecting the nodulation competitiveness. In a peat substrate, *S. meliloti* strains can be “hitch-hiked” by nematodes and being transported on their surface or into their gut (Horiuchi et al., 2005). Percolating water and tillage, may help the vertical motility of rhizobia in soil and long the growing roots (López-García et al., 2002). In *B. japonicum* the influence of motility on competitiveness may deeply depend on the water status of the rooting substrate (Althabegoiti et al., 2001) or the soil exploration by growing roots (Liu et al., 1989). Traits related to drought tolerance might also be relevant in the rhizosphere and during the root infection. Indeed, facing dry environments, one of the common strategies is the accumulation of solutes as trehalose. Without affecting its intrinsic nodulation and N_2 fixation efficiency, *S. meliloti* triple mutant (*treS*, *treY*, and *otsA*) unable to accumulate trehalose showed compromised competitive abilities (Domínguez-Ferreras et al., 2009). Moreover, as observed in *S. meliloti*, greater ability in accumulating and metabolizing trehalose improved the penetration of root hairs (Ampomah et al., 2008).

Forms of indirect (or exploitative) competition can also be critical in the rhizosphere and for the subsequently root colonization (Checcucci et al., 2017b). Some examples are the ability of rhizobia in efficiently metabolizing carbon sources, (Gage and Long, 1998; Ramachandran et al., 2011) specifically glycerol in *R. leguminosarum* (Ding et al., 2012), but also nitrogen sources as proline in *S. meliloti* (van Dillewijn et al., 2001) from root exudates. Also precluding competitor’s activity may be a relevant skill in the rhizosphere, for example through the overproduction of siderophores (diCenzo et al., 2014) or the production of antibacterial compounds as bacteriocin. In *R. etli* the introduction of genes encoding for the bacteriocin trifolitoxin, originally found in *R. leguminosarum* bv *trifolii*, significantly improved the competition for nodulation in agricultural settings (Robleto et al., 1998). Moreover, bacteriocin-encoding genes cluster can be found within genomes of members of the genus *Sinorhizobium*, suggesting a possible role in inter and intra-species dynamics competition (Checcucci et al., 2017b).

Beyond the rhizosphere colonization, rhizobial adhesion to root surfaces is also crucial for successful nodulation. For instance, mutants of *B. japonicum* unable to incorporate galactose in their exopolysaccharide (EPS), required for the root hair invasion, are severely impaired in the plant-lectin binding, showing a lowered adhesion to soybean roots and infective capacity (Pérez-Giménez et al., 2009; Quelas et al., 2006; 2010). Moreover, the over-expression of two regulatory genes *pssA* and *rosR* for EPS production increased competition abilities for clover nodulation in *R. leguminosarum* bv *trifolii* and *R. etli* (Bittinger et al., 1997; Janczarek et al., 2009). Attachment and aggregation by rhizobia to roots involv a group of rhizobial adhering

proteins (specifically calcium-binding lectins Raps) (Ausmees et al., 2001). In *R. leguminosarum* bv *trifolii* the overproduction of the bacterial adhesion protein RapA1 have a positive effect on competition for nodule occupation increasing rhizobial adhesion to roots (Mongiardini et al., 2009), and Rap-mediated interactions between rhizobial cells may improve the accumulation on root hairs and thus competitiveness (Frederix et al., 2014). The role of biofilm formation is also crucial for an optimal root colonization and symbiosis between *S. fredii* and *G. max* (Pérez-Montano et al., 2014).

The plant growth regulator ethylene is also an inhibitor of nodulation. Strains able to locally counteract the nodulation inhibition by ethylene seem more competitive for nodulation. Hence, when the nodulation process is underway, the ability to modulate the plant ethylene levels by rhizobia strains, and thus to prevent nodulation inhibition by the plant, can also be interpreted as an enhanced competitive trait. In *B. elkanii*, a higher rhizobitoxine production, which is an ethylene synthesis inhibitor, results in greater nodulation competitiveness on *Macroptilium atropurpureum* (Siratro). (Okazaki et al., 2003). Moreover, the enzyme 1-aminocyclopropane-1-carboxylate (ACC) deaminase (AcdS), involved in the cleaving of the precursor of ethylene ACC, seems to play an important role in lowering the plant-produced ethylene levels during the rhizopheric colonization (Checcucci et al., 2017) or during initial infection (Conforte et al., 2010).

1.2.2 Methodological approaches to study the competition for nodulation

The employment of co-inoculation of different strains on the same plant individual is a common practice of the competition studies reported so far. In methodological approaches for the study of competition, the general criterion is to apply different labeling of the competitor strains to discriminate the nodules occupied singly by the two strains, or to ascertain the presence of mixed nodules both (Pérez-Giménez et al., 2011). The methods employed to strains labelling are not different than those used for strain selection or tracking and including: antibiotic resistances (Iturralde et al., 2019), fluorescent proteins (Gage et al., 1996; Gage et al., 2002; Checcucci et al., 2016; Lardi et al., 2017; Mendoza-Suárez et al., 2020), or reporter proteins

such as galactosidase or β-glucuronidase (GUS) (Okazaki et al., 2003; Ding et al., 2012; Frederix et al., 2014).

The phenotype features mentioned above (Chapter 1.2.1), and probably related to the competition abilities, partially shed a light on this highly complex phenotype. However, to study the function of genes related to the competition only single-gene approaches have been employed so far, as gene disruption or deletion, usually followed by a comparison of competition phenotype between the obtained mutants and their wild-type. What is currently lacking is a "whole" comprehension of the genetic basis of this phenomenon. To elucidate the total genes associated with competitiveness with omics-approaches, few preliminary studies were conducted. An early attempt was conducted by Pobigaylo and collaborators (2008). To identify the "whole" genetic determinants involved in symbiosis and competitiveness, a large library (378) of signature-tagged mutants (STM) of *S. meliloti* Rm2011 was generated and tested on *M. sativa*. Among the STM, 23 were altered in genes that strongly affected the individual competition assays, which are related to metabolic function (synthesis of tryptophan and EPS, myo-inositol catabolism), transport functions (of metal ion, amino acid, and export of capsular polysaccharides), and transcriptional regulation (expression of transcriptional repressor and GTP-binding elongation factor). Similarly, STM libraries were also developed for genome-wide identification of several genes involved in symbiotic competitiveness in *M. loti* (Borjigin et al. 2011) and in *S. fredii* (Wang et al. 2016). A transcriptomic approach has been employed to point out genes involved in the competitive nodulation process. In particular, transcriptomic analyses were performed to investigate the effect of soybean root exudates on two *B. diazoefficiens* strains with different competition abilities (Liu et al., 2017). In this study the identified genetic determinants underlying rhizobial competitiveness partially overlap with those reported above. Hence, functions associated with metabolic fitness [proline and indole-3-acetic acid (IAA) metabolism], transport of amino acid and carbon sources by ABC-transporter proteins), bacterial chemotaxis, and related to "two-component regulation system" may be responsible for higher competition abilities (Liu et al., 2017). Also, general adaptations to legumes rhizosphere and specific adaptations in pea rhizosphere were enlightened in detail by comparing *R. leguminosarum* colonization of the rhizosphere of its host legume with that of a non-host legume (Ramachandran et al., 2011).

1.2.3 The competition problem in agricultural settings

In agricultural settings, the competition dynamics are known to be responsible for the so-called "rhizobial competition problem" (Yates et al., 2011). Usually, the soils are populated by a consistent amount of local indigenous rhizobia that are well adapted to the rhizosphere and, in general, to soil conditions (Sadowsky and Graham, 1998). Despite these resident rhizobia are numerically dominant, they may be not enough efficient in N_2-fixing once established the symbiosis, especially with introduced legumes (as soybean in European soil) (Yates et al., 2011; Iturralde et al., 2019). More important, as a consequence of several years of adaptation, the rhizobial populations were enriched in genotypes with enhanced competitiveness and strong abilities to persist in the local environment, allowing the appearance of "cheaters" that weakly fixed N_2 (Iturralde et al., 2019). Accordingly, the competition abilities for plant nodulation of soil rhizobia population may harshly challenge the efficiency of rhizobial inoculants to occupy a significant proportion of the nodules (Triplett and Sadowsky, 1992; Toro, 1996; Howieson and Ballard, 2004). Besides, it has to be considered that the physiological state of the rhizobia potentially employed as bio-inoculant is different from the corresponding resident rhizobia. Indeed, the inoculant strains are often well-adapted to laboratory conditions when they were frequently grown in rich broths, then likely unable to face stress conditions that may encounter once released in soil. Hence, because of the competition problem, the N_2-fixation process and the yield of the legume crops following the introduced inoculant can be compromised and resulted in substantial economic consequences (Yates et al., 2011). Therefore, a deepened knowledge of traits related to competition is of paramount importance to improve the highly N_2 fixation effective inoculant strains also in terms of competition performance.

1.3 The quest for elite inoculant strains: integrated genomics approaches

The ecological and economical importance of the legume-rhizobia interaction has encouraged the production of commercial rhizobia inoculants (Checcucci et al., 2017b). It is well known that in sustainable agriculture, a common approach is to inoculate legume crops with compatible rhizobia that fix substantial amounts of nitrogen, mainly selected based on experiments that

showed increased legume crops. However, the longstanding competition problem, described in the previous chapter, harshly challenges the feasibility of rhizobial bio-inoculants, limiting or completely counteracting their positive effects on crop yield (Dwivedi et al., 2015). Consequently, in the last years, the awareness that the social aspects of interactions among rhizobia strains (and among inoculant strains with native rhizosphere populations) should be considered in the development of rhizobia inoculants (Checcucci et al., 2017b) is arising. In particular, it is crucial to recognize, and improve, the competitiveness features of the strains as an additional criterion beyond the high N_2- fixation efficiency (Lupwayi et al., 2006; Schumpp and Deakin, 2010; Irisarri et al., 2019) for the selection of successful rhizobia-based "elite" inoculants. However, the identification of this "elite" strains that are simultaneously highly effective and competitive can be a time-consuming process, which is primarily based on plentiful studies in laboratory conditions and subsequently on the assessment of the selected strains in greenhouse and field trials (Bourion et al., 2018, Irisarri et al., 2019). To accelerate the screening process, a reporter plasmid-based library was developed for the concurrent evaluation of i) measurement of N_2-fixation in individual nodules, using a consensus *nifH* promoter (*PsnifH*) that drive the expression of the reported gene for green fluorescent protein (GFP), and ii) the rhizobial competitiveness, with barcode strain identification (Plasmid ID) through a next-generation sequencing (NGS) approach (Mendoza-Suárez et al., 2020).

The last ten years were witnessed of increase in pioneering synthetic biology approaches that address the challenge of improving significant phenotypes for further "elite" rhizobial inoculants in sustainable agriculture. The great phylogenetic diversity that marked out the paraphyletic group of rhizobia, together with the feasibility of large pangenomes (i.e., the gene set current in a group of microbial strains belonging to the same species) (Medini et al., 2005; Galardini et al., 2015b; Young et al., 2016) and multipartite genomes (Galibert et al., 2001; Galardini et al., 2013), foresee the possibility of large-scale genome manipulation approaches in which entire replicons may be combined in a single genome. A *S. meliloti* cis-hybrid strain was indeed produced by the transfer of the symbiotic megaplasmid pSymA (that harbor gene related to the competition, symbiotic capabilities, nitrogen fixation process and tolerance to harsh soils) from a donor strain to a closely related *S. meliloti* strain (Checcucci et al., 2018). The resulted cis-hybrid strain displayed a host cultivar-specific improvement in symbiotic effectiveness inferring that plasmid transfer may be a practicable method for "breeding" boosted strains inoculants. Interestingly, experimental evolution approaches under lab conditions were also used to mimic the initial LGT (lateral gene transfer) event by transferring the symbiotic plasmid of *C. taiwanensis* into the pathogen *Ralstonia solanacearum*,

demonstrating the possibility of converting non-rhizobia into strains that prompt the nodule formation and intracellular infection on atypical hosts (Marchetti et al., 2010; Guan et al., 2013).

The link between a phenotype and its genetic basis is one of the greatest challenges of biology. Besides forward and reverse genetics approaches, genotype-phenotype association analyses, or genome-wide association (GWAS) studies, are a complementary approach for the identification of gene function (Epstein et al., 2018). By using statistical approaches, GWAS analysis effectively associates naturally occurring allelic variation with specific phenotypic traits (Epstein et al., 2018). This powerful approach was largely employed for identifying genes responsible for quantitative trait variation in plants and animals (Curtin et al., 2017). However, in the last years, the GWAS analysis was successfully applied in bacteria (Chen & Shapiro, 2017) to reveal the genetic basis of phenotypes as antibiotic resistance (Alam et al., 2014), virulence (Maury et al., 2016), and bacterial host preference (Sheppard et al., 2013). In particular, using GWAS analysis, candidate genes associated with several phenotypes were picked in rhizobia, as novel operons responsible for nickel tolerance in *Mesorhizobium* strain (Porter et al., 2017) or as candidate alleles and gene presence-absence variants underlying symbiosis traits, antibiotic resistance, and use of various carbon sources in *S. meliloti* (Epstein et al., 2018). More importantly, the marker genes controlling the host-specific competitiveness for nodulation in *R. leguminosarum* symbiovar *viciae* (Rlv) were also investigated (Boivin et al., 2020). In this work, by using GWAS analysis, in addition to *nod* genes other genomic regions, mainly located on Rlv accessory plasmids and thus potentially horizontally transferred, were associated to the competition phenotype (Boivin et al., 2020). Therefore, GWAS analysis could be a powerful tool to successfully pinpoint the genetic determinants responsible for competition-related phenotypes. Accordingly, GWAS and similar approaches may be useful in designing *ad hoc* inoculant strains able to fit with desired legume crops and general soil conditions, to overcome ‘the competition problem’ detected in agricultural settings.

1.4 References

Abdian, P.L., Caramelo, J.J., Ausmees, N., Zorreguieta, A. (2013). RapA2 is a calcium-binding lectin composed of two highly conserved cadherin-like domains that specifically recognize *Rhizobium leguminosarum* acidic exopolysaccharides. J. Biol. Chem., 288: 2893-2904.

Alam, M.T., Petit, R.A., Crispell, E.K., Thornton, T.A., Conneely, K.N., Jiang, Y., Satola, S.W., Read, T.D. (2014). Dissecting vancomycin-intermediate resistance in *Staphylococcus aureus* using genome-wide association. Genome Biol. Evol., 6: 1174-1185.

Allaway, D., Lodwig, E.M., Crompton, L.A., Wood, M., Parsons, R., Wheeler, T.R., Poole, P.S. (2000). Identification of alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea bacteroids. Mol. Microbiol., 36(2): 508-515.

Althabegoiti, M.J., Covelli, J.M., Pérez-Giménez, J., Quelas, J.I., Mongiardini, E.J., López, M.F., López-García, S.L., Lodeiro, A.R. (2011). Analysis of the role of the two flagella of *Bradyrhizobium japonicum* in competition for nodulation of soybean. FEMS Microbiol. Lett., 319(2): 133-139.

Amarger, N. (1981). Competition for nodule formation between eff ective and ineffective strains of *Rhizobium meliloti*. Soil Biol. Biochem., 13: 475-480.

Ampomah, O.Y., Beck-Jensen, J., Bhuvaneswari, T.V. (2008). Lack of trehalose catabolism in *Sinorhizobium* species increases their nodulation competitiveness on certain host genotypes. New Phytol., 179: 495-504.

Andrews, M., Andrews, M.E. (2017). Specificity in legume-rhizobia symbioses. Int. J. Mol. Sci., 18:705.

Ausmees, N., Jacobsson, K., Lindberg, M. (2001). A unipolarly located, cell-surface-associated agglutinin RapA belongs to a family of *Rhizobium*-adhering proteins (Rap) in *Rhizobium leguminosarum* bv. *trifolii*. Microbiol., 147: 549-559.

Barnett, M.J, Fisher, R.F., (2006). Global gene expression in the rhizobial–legume symbiosis. Symbiosis, 42: 1-24.

Bascompte, J., Jordano, P. (2007). Plant-animal mutualistic networks: the architecture of biodiversity. Annu. Rev. Ecol. Evol. Syst., 38: 567-593.

Becker, A., Fraysse, N., Sharypova, L. (2005). Recent advances in studies on structure and symbiosis-related function of rhizobial K-antigens and lipopolysaccharides. Mol. Plant Microbe Interact., 18: 899-905.

Bellabarba, A., Fagorzi, C., diCenzo, G.C., Pini, F., Viti, C., Checcucci, A. (2019). Deciphering the symbiotic plant microbiome: translating the most recent discoveries on rhizobia for the improvement of agricultural practices in metal-contaminated and high saline lands. Agronomy, 9:529.

Bergersen, F.J., Peoples, M.B., Turner, G.L. (1991). A role for poly-b-hydroxybutyrate in bacteroids of soybean nodules. Proc. R. Soc. Lond. B., 245: 59-64.

Bittinger, M.A., Milner, J.L., Saville, B.J., Handelsman, J. (1997). *rosR*, a determinant of nodulation competitiveness in *Rhizobium etli*. Mol. Plant Microbe Interact., 10: 180-186.

Boivin, S., Lahmidi, N.A., Sherlock, D., Bonhomme, M., Dijon, D., Heulin-Gotty, K., Le-Queré, A., Pervent, M., Tauzin, M., Carlsson, G., Jensen, E., Journet, E.P., Lopez-Bellido, R., Seidenglanz, M., Marinkovic, J., Colella, S., Brunel, B., Young, P., Lepetit, M. (2020). Host-specific competitiveness to form nodules in *Rhizobium leguminosarum* symbiovar viciae. New Phytol., 226: 555-568.

Borjigin, N., Furukawa, K., Shimoda, Y., Tabata, S., Sato, S., Eda, S., Minamisawa, K., Mitsui, H. (2011). Identification of *Mesorhizobium loti* genes relevant to symbiosis by using signature-tagged mutants. Microbes Environ., 26(2): 165-171.

Bourion, V., Heulin-Gotty, K., Aubert, V., Tisseyre, P., Chabert-Martinello, M., Pervent, M., Delaitre, C., Vile, D., Siol, M., Duc, G., Brunel, B., Burstin, J., Lepetit, M. (2018). Co-inoculation of a pea core-collection with diverse rhizobial strains shows competitiveness for nodulation and efficiency of nitrogen fixation are distinct traits in the interaction. Front. Plant Sci., 8:2249

Brewin, N.J. (2004). Plant cell wall remodelling in the *Rhizobium*–Legume symbiosis. Crit. Rev. Plant Sci., 23: 293-316.

Brockwell, J., Schwinghamer, E. A., Gault, R. R. (1977). Ecological studies of root-nodule bacteria introduced into field environments —V.A critical examination of stability of antigenic and streptomycin-resistance markers for identification of strains of *Rhizobium trifolii*. Soil Biol. Biochem., 9: 19-24.

Bronstein, J.L. (2009). The evolution of facilitation and mutualism. J. Ecol., 97: 1160-1170.

Burdon, J.J., Gibson, A.H., Searle, S.D., Woods, M.J, Brockwell, J. (1999). Variation in the effectiveness of symbiotic associations between native rhizobia and temperate Australian *Acacia*: within-species interactions. J. Appl. Ecol., 36: 398-408.

Caetano-Anollés, G., Wall, L.G., De Micheli, A.T., Macchi, E M, Bauer, W.D, Favelukes, G. (1988b). Role of motility and chemotaxis in efficiency of nodulation by *Rhizobium meliloti*. Plant Physiol., 86: 1228-1235.

Capela, D., Carrere, S., Batut, J. (2005). Transcriptome-based identification of the *Sinorhizobium meliloti* NodD1 regulon. Appl. Environ. Microbiol., 71: 4910-13

Cevallos, M.A., Encarnacion, S., Leija, A., Mora, Y., Mora, J. (1996). Genetic and physiological characterization of a *Rhizobium etli* mutant strain unable to synthesize poly-b-hydroxybutyrate. J. Bacteriol., 178: 1646-1654.

Checcucci, A., Azzarello, E., Bazzicalupo, M., De Carlo, A., Emiliani, G., Mancuso, S., Spini, G., Viti, C., Mengoni, A. (2017a). Role and regulation of ACC deaminase gene in *Sinorhizobium meliloti*: is it a symbiotic, rhizospheric or endophytic gene? Front. Genet., 8:6.

Checcucci, A., Azzarello, E., Bazzicalupo, M., Galardini, M., Lagomarsino, A., Mancuso, S., Marti, L., Marzano, M.C., Mocali, S., Squartini, A., Zanardo, M., Mengoni, A. (2016). Mixed nodule infection in *Sinorhizobium meliloti–Medicago sativa* symbiosis suggests the presence of cheating behavior. Front. Plant Sci., 7: 835.

Checcucci, A., diCenzo, G.C., Bazzicalupo, M., Mengoni, A. (2017). Trade, diplomacy, and warfare: the quest for elite rhizobia inoculant strains. Front. Microbiol., 8:2207.

Checcucci, A., diCenzo, G.C., Ghini, V., Bazzicalupo, M., Becker, A., Decorosi, F., Döhlemann, J., Fagorzi, C., Finan, T.M., Fondi, M., Luchinat, C., Turano, P., Vignolini, T., Viti, C., Mengoni A. (2018). Creation and Characterization of a Genomically Hybrid Strain in the Nitrogen-Fixing Symbiotic Bacterium *Sinorhizobium meliloti*. ACS Synth. Biol., 7: 2365-2378.

Chen, P.E., Shapiro, B.J. (2015). The advent of genome-wide association studies for bacteria. Curr. Opin. Microbiol., 25: 17-24.

Conforte, V.P., Echeverría, M., Sánchez, C., Ugalde, R.A., Menéndez, A.B., Lepek, V.C. (2010). Engineered ACC deaminase-expressing free-living cells of *Mesorhizobium loti* show increased nodulation efficiency and competitiveness on *Lotus* spp. J. Gen. Appl. Microbiol., 56: 331-338.

Curtin, S.J., Tiffin, P., Guhlin, J., Trujillo, D.I., Burghart, L.T., Atkins, P., Baltes, N.J., Denny, R., Voytas, D.F., Stupar, R.M., Young, N.D. (2017). Validating genome-wide association candidates controlling quantitative variation in nodulation. Plant Physiol., 173: 921-931.

Daubech, B., Remigi, P., Doin de Moura, G., Marchetti, M., Pouzet, C., Auriac, M.C., Gokhale, C.S., Masson-Boivin, C., Capela, D. (2017). Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation. eLife, 6: 1-21.

De Mita, S. (2012). For better or for worse: cooperation and competition in the legume–rhizobium symbiosis. New Phytol., 194: 885-887.

Denison, R.F. (2000). Legume sanctions and the evolution of symbiotic cooperation by rhizobia. American Naturalist., 156(6): 567-576.

Denison, R.F., Bledsoe, C., Kahn, M., O'Gara, F., Simms, E.L., Thomashow, L.S. (2003a). Cooperation in the rhizosphere and the "free rider" problem. Ecol., 84: 838-45

Denison, R.F., Kiers, E.T. (2004). Lifestyle alternatives for rhizobia: mutualism, parasitism, and forgoing symbiosis. Fems Microbiol., 237: 87-193.

Denison, R.F., Kiers, E.T. (2011). Life histories of symbiotic rhizobia and mycorrhizal fungi. Curr. Biol., 21: R775-R785.

diCenzo, G. C., Maclean, A. M., Milunovic, B., Golding, G. B., Finan, T. M. (2014). Examination of prokaryotic multipartite genome evolution through experimental genome reduction. PLOS Genet. 10: e1004742.

diCenzo, G.C., Checcucci, A., Bazzicalupo, M., Mengoni, A., Viti, C., Dziewit, L., Finan, T.M., Galardini, M., Fondi, M. (2016). Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. Nat. Commun., 7: 1-10.

diCenzo, G.C., Tesi, M., Pfau, T., Mengoni, A., Fondi, M. (2020). Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. Nat. Commun., 11: 1-11.

Ding, H., Yip, C. B., Geddes, B. A., Oresnik, I. J., Hynes, M. F. (2012). Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. Microbiology, 158: 1369-1378.

Dixon, R., Wheeler, C.T. (1986). Nitrogen fixation in plants. Blackie and Son, Glascow, United Kingdom.

Domínguez-Ferreras, A., Soto, M.J., Pérez-Arnedo, R., Olivares, J., Sanjuán J. (2009). Importance of trehalose biosynthesis for *Sinorhizobium meliloti* osmotolerance and nodulation of alfalfa roots. J. Bacteriol., 191: 7490-7499.

Downie, J.A. (2010). The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. FEMS Microbiol. Rev., 34: 150-170.

Dwivedi, S. L., Sahrawat, K. L., Upadhyaya, H. D., Mengoni, A., Galardini, M., Bazzicalupo, M., Biondi, E.G., Hungria, M., Kaschuk, G., Blair, M.W., Ortiz, R. (2015). Advances in host plant and rhizobium genomics to enhance symbiotic nitrogen fixation in grain legumes. Adv. Agron., 129: 1-116.

Epstein, B., Abou-Shanab, R.A.I., Shamseldin, A., Taylor, M.R., Guhlin, J., Burghardt, L.T., Nelson, M., Sadowsky, M.J., Tiffin, P. (2018). Genome-wide association analyses in the model rhizobium *Ensifer meliloti*. mSphere, 3: e00386-18.

Fischer, H.M. (1994). Genetic regulation of nitrogen fixation in rhizobia. Microbiol. Rev., 58: 352-386.

Franssen, H.J., Vijn, I., Yang, W.C., Bisseling, T. (1992). Developmental aspects of the *Rhizobium*-legume symbiosis. Plant Mol Biol., 19: 89-107.

Frederickson, M.E. (2013). Rethinking mutualism stability: cheaters and the evolution of sanctions. Q. Rev. Biol., 88: 269-295.

Frederix, M., Edwards, A., Swiderska, A., Stanger, A., Karunakaran, R., Williams, A., Abbruscato, P., Sanchez-Contreras, M., Poole, P.S., Downie, J.A. (2014). Mutation of *praR* in *Rhizobium leguminosarum* enhances root biofilms, improving nodulation competitiveness by increased expression of attachment proteins. Mol. Microbiol., 93: 464-478.

Friesen, M.L., Heath, K.D. (2013). One hundred years of solitude: integrating single-strain inoculations with community perspectives in the legume-rhizobium symbiosis. New Phytol., 198: 7-9.

Friesen, M. L., Mathias, A. (2010). Mixed infections may promote diversification of mutualistic symbionts: why are there ineffective rhizobia? J. Evol. Biol., 23: 323-334.

Friesen, M.L. (2012). Widespread fitness alignment in the legume–rhizobium symbiosis. New Phytol., 194: 1096-1111

Gage, D.J. (2002). Analysis of infection thread development using GFP- and DsRed-expressing *Sinorhizobium meliloti*. J. Bacteriol., 184: 7042-7046.

Gage, D.J. (2002). Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*. J. Bacteriol., 184: 7042-7046.

Gage, D.J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol. Mol. Biol. Rev., 68: 280-300.

Gage, D.J., Bobo, T., Long, S.R., (1996). Use of green fluorescent protein to visualize the early events of symbiosis between *Rhizobium meliloti* and alfalfa (*Medicago sativa*). J. Bacteriol., 178: 7159-7166.

Gage, D.J., Long, S.R (1998). α-Galactoside uptake in *Rhizobium meliloti*: isolation and characterization of *agpA*, a gene encoding a periplasmic binding protein required for melibiose and raffinose utilization. J. Bacteriol., 180: 5739-5748.

Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G., Mengoni, A. (2013). Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*. Genome Biol. Evol., 5: 542-558.

Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E.G., and Mengoni, A. (2015b). "Pangenome evolution in the symbiotic nitrogen fixer *Sinorhizobium meliloti*" in Biological Nitrogen Fixation, (ed.) F.J. de Bruijn (New York, NY: John Wiley & Sons, Inc), 265–270.

Galibert, F., Finan, T. M., Long, S. R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M.J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R.W., Dreano, S., Federspiel, N.A., Fisher, R.F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R.W., Jones, T., Kahn, D., Kahn, M.L., Kalman, S., Keating, D.H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M.C., Pohl, T.M., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thebault, P., Vandenbol, M., Vorhölter, F.J., Weidner, S., Wells, D.H., Wong, K., Yeh, K.C., Batut, J. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science, 293: 668-672.

Gibson, K.E., Kobayashi, H., Walker, G.C. (2008). Molecular determinants of a symbiotic chronic infection. Annu. Rev. Genet., 42: 413-41.

González, E.M., Aparicio-Tejo, P.M., Gordon, A.J., Minchin, F.R., Royuela, M., Arrese-Igor, C. (1998). Water-deficit effects on carbon and nitrogen metabolism of pea nodules. J. Exp. Bot., 49: 1705-1714.

Guan, S.H., Gris, C., Cruveiller, S., Pouzet, C., Tasse, L., Leru, A., Maillard, A., Medigue, C., Batut, J., Masson-Boivin, C., Delphine, C. (2013). Experimental evolution of nodule intracellular infection in legume symbionts. ISME J., 7: 1367-1377.

Gubry-Rangin, C., Garcia, M., Béna, G. (2010). Partner choice in *Medicago truncatula-Sinorhizobium* symbiosis. Proc. Biol. Sci., 277: 1947-51.

Hahn, M., Studer, D. (1986). Competitiveness of a nif- *Bradyrhizobium japonicum* mutant against the wild-type strain. FEMS Microbiol. Lett., 33: 143-148.

Hakoyama, T., Niimi, K., Watanabe, H., Tabata, R., Matsubara, J., Sato, S., Nakamura, Y., Tabata, S., Jichun, L., Matsumoto, T. (2009). Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. Nature, 462: 514-518.

Hardin, G. (1968). The tragedy of the commons. Science, 162: 1243-48

Heath, K.D, Tiffin, P. (2007). Context dependence in the coevolution of plant and rhizobial mutualists. Proc. R. Soc. Lond. B., 274: 1905-1912.

Heath, K.D., Tiffin, P. (2009). Stabilizing mechanisms in a legume–rhizobium mutualism. Evolution., 63: 652-662.

Herridge, D.F., Peoples, M.B., Boddey, R.M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil, 311: 1-18.

Holland, J.N., Ness, J.H., Boyle, A., Bronstein, J.L. (2005). "Mutualisms as consumer-resource interactions", in Ecology of Predator-Prey Interactions, (ed) Barbosa P. (Oxford University Press), 17-33.

Horiuchi, J., Prithiviraj, B., Bais, H.P., Kimball, B.A., Vivanco, J.M. (2005). Soil nematodes mediate positive interactions between legume plants and rhizobium bacteria. Planta, 222: 848-857.

Howieson, J.G., Ballard, R. (2004). Optimising the legume symbiosis in stressful and competitive environments within southern Australia – some contemporary thoughts. Soil Biol. Biochem., 36: 1261-1273.

Irisarri, P., Cardozo, G., Tartaglia, C., Reyno, R., Gutiérrez, P., Lattanzi, F. A., Rebuffo, M., Monza, J. (2019). Selection of competitive and efficient rhizobia strains for white clover. Front. Microbiol., 10: 1-11.

Iturralde, E.T., Covelli, J.M., Alvarez, F., Perez-Gimenez, J., Arrese-Igor, C., Lodeiro, A.R. (2019). Soybean-nodulating strains with low intrinsic competitiveness for nodulation, good symbiotic performance, and stress tolerance isolated from soybean-cropped soils in Argentina. Front. Microbiol., 10:1061.

Janczarek, M., Jaroszuk-Sciseł, J., Skorupska, A. (2009). Multiple copies of *rosR* and *pssA* genes enhance exopolysaccharide production, symbiotic competitiveness and clover nodulation in *Rhizobium leguminosarum* bv. *trifolii*. Ant. van Leeuw., 96: 471-486.

Jones, E.I., Afkhami, M.E., Akcay, E., Bronstein, J.L., Bshary, R., Frederickson, M.E., Heath, K.D., Hoeksema, J.D., Ness, J.H., Pankey, M.S., Porter, S.S., Sachs J,L., Scharnagl. K., Friesen, M,L. (2015). Cheaters must prosper: reconciling theoretical and empirical perspectives on cheating in mutualism. Ecol. Lett., 18: 1270-1284

Jones, E.I., Bronstein, J.L., Ferriere, R. (2012). The fundamental role of competition in the ecology and evolution of mutualisms. Ann. N. Y. Acad. Sci., 1256: 66-88.

Jones, K.M., Kobayashi, H., Davies, B.W., Taga, M.E., Walker, G.C. (2007). How rhizobial symbionts invade plants: The *Sinorhizobium-Medicago* model. Nat. Rev. Microbiol., 5: 619-633.

Kaiser, B.N., Moreau, S., Castelli, J., Thomson, R., Lambert, A., Bogliolo, S., Puppo, A., Day, D.A. (2003). The soybean NRAMP homologue, GmDMT1, is a symbiotic divalent metal transporter capable of ferrous iron transport. Plant J., 35: 295-304.

Karunakaran, R., Ramachandran, V.K., Seaman, J.C., East, A.K., Moushine, B., Mauchline, T. H., Prell, J., Skeffington, A. Poole, P.S. (2009). Transcriptomic analysis of *Rhizobium leguminosarum* bv. *viciae* in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. J. Bacteriol., 191: 4002-4014.

Kiers, E. T., Ratcliff, W. C., Denison, R. F. (2013). Single-strain inoculation may create spurious correlations between legume fitness and rhizobial fitness. New Phytol., 198: 4-6.

Kiers, E.T., Rousseau, R. A., Denison, R. F. (2006). Measured sanctions: Legume hosts detect quantitative variation in rhizobium cooperation and punish accordingly. Evol. Ecol. Res., 8: 1077-1086.

Kiers, E.T., Denison, R.F. (2008). Sanctions, cooperation, and the stability of plant–rhizosphere mutualisms. Annu. Rev. Ecol. Evol. Syst., 39: 215-236.

Kiers, E.T., Hutton, M.G., Denison, R.F. (2007). Human selection and the relaxation of legume defences against ineffective rhizobia. Proc. R. Soc. B., 274: 3119-3126.

Kiers, E.T., Palmer, T.M., Ives, A.R., Bruno, J.F. & Bronstein, J.L. (2010). Mutualisms in a changing world: an evolutionary perspective. Ecol. Lett., 13: 1459-1474.

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

Kosslak, R.M., Bohlool, B.B., Dowdle, S., Sadowsky, M.J. (1983). Competition of *Rhizobium japonicum* strains in early stages of soybean nodulation. Appl. Environ. Microbiol., 46: 870-873.

Krusell, L., Krause, K., Ott, T., Desbrosses, G., Kramer, U., Sato, S., Nakamura, Y., Tabata, S., James, E.K., Sandal, N., Stougaard, J., Kawaguchi, M., Miyamoto, A., Suganuma, N., Udvardi, M.K. (2005). The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. Plant Cell., 17: 1625-36.

Lagarda-Diaz, I., Guzman-Partida, A.M., Vazquez-Moreno, L. (2017). Legume lectins: proteins with diverse applications. Int. J. Mol. Sci., 18(6): 1242.

Laguerre, G., Bardin, M., Amarger, N. (1993). Isolation from soil of symbiotic and nonsymbiotic *R. leguminosarum* by DNA hybridization. Can. J. Microbiol., 39: 1142-1149.

Lardi, M., de Campos, S.B., Purtschert, G., Eberl, L., Pessi, G. (2017). Competition experiments for legume infection identify *Burkholderia phymatum* as a highly competitive β-rhizobium. Front. Microbiol., 8:1527.

Laus, M.C., van Brussel, A.A., Kijne, J.W. (2005). Role of cellulose fibrils and exopolysaccharides of *Rhizobium leguminosarum* in attachment to and infection of *Vicia sativa* root hairs. Mol. Plant. Microbe Interact., 18: 533-538.

Leung, T.L., Poulin, R. (2008). Parasitism, commensalism and mutualism: exploring the many shades of symbioses. Vie Milieu-Life Environ., 58: 107-115.

Ling, J., Zheng, H., Katzianer, D.S., Wang, H., Zhong, Z., Zhu J., (2013). Applying reversible mutations of nodulation and nitrogen-fixation genes to study social cheating in *Rhizobium etli*-legume interaction. PLoS One, 8(7): e70138.

Liu, R., Tran, V.M., Schmidt, E.L. (1989), Nodulating competitiveness of a nonmotile Tn7 mutant of *Bradyrhizobium japonicum* in nonsterile soil. Appl. Environ. Microbiol., 55: 1895-1900.

Liu, Y., Jiang, X., Guan, D., Zhou, W., Ma, M., Zhao, B., Cao, F., Li, L., Li, J. (2017). Transcriptional analysis of genes involved in competitive nodulation in *Bradyrhizobium diazoefficiens* at the presence of soybean root exudates. Sci. Rep., 7:10946.

Lodeiro, A.R., Lopez-Garcia, S.L., Vazquez, T.E., Favelukes, G. (2000). Stimulation of adhesiveness, infectivity, and competitiveness for nodulation of *Bradyrhizobium japonicum* by its pretreatment with soybean seed lectin. FEMS Microbiol. Lett., 188: 177-84.

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

Lodwig, E.M., Poole P.S. (2003). Metabolism of *Rhizobium* bacteroids. Crit. Rev. Plant Sci., 22: 37-78.

López-García, S.L., Vázquez, T.E.E., Favelukes, G., Lodeiro, A.R. (2002). Rhizobial position as a main determinant in the problem of competition for nodulation in soybean. Environ. Microbiol., 4: 216-224.

Lupwayi, N. Z., Clayton, G. W., Rice, W. A. (2006). Rhizobial inoculants for legume crops. J. Crop Improv., 15: 289-321.

Marchetti, M., Capela, D., Glew, M., Cruveiller, S., Chane-Woon-Ming, B., Gris, C., Timmers, T., Poinsot, V., Gilbert, L.B., Heeb, P., Médigue, C., Batut, J., Masson-Boivin, C. (2010). Experimental evolution of a plant pathogen into a legume symbiont. PLoS Biol., 8: e1000280.

Marco, D.E., Carbajal, J.P., Cannas, S., Pérez-Arnedo, R., Hildalgo-Perea, A., Olivares, J., Ruiz-Sainz, J. E., Sanjuán, J. (2009). An experimental and modelling exploration of the host-sanction hypothesis in legume–rhizobia mutualism. J. Theor. Biol., 259: 423-433.

Masson-Boivin, C., Giraud, E., Perret, X., Batut, J. (2009). Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? Trends Microbiol., 17: 458-466.

Masson-Boivin, C., Sachs, J.L. (2018). Symbiotic nitrogen fixation by rhizobia-the roots of a success story. Curr. Opin. Plant Biol., 44: 7-15.

Matamoros, M.A., Baird, L.M., Escuredo, P.R., Dalton, D.A., Minchin, F.R., Iturbe-Ormaetxe, I., Rubio, M.C., Moran, J.F., Gordon, A.J., Becana, M. (1999). Stress-induced legume root nodule senescence: physiological, biochemical, and structural alterations. Plant Physiol., 121: 97-111.

Maury, M.M., Tsai, Y.H., Charlier, C., Touchon, M., Chenal-Francisque, V., Leclercq, A., Criscuolo, A., Gaultier, C., Roussel, S., Brisabois, A., Disson, O., Rocha, E.P.C., Brisse, S., Lecuit, M. (2016). Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. Nat. Genet., 48: 308-313.

Medini, D., Donati, C., Tettelin, H., Masignani, V., Rappuoli, R. (2005). The microbial pan-genome. Curr. Opin. Genet. Dev., 15: 589-594.

Mellor, H.Y., Glenn, A.R., Arwas, R. Dilworth, M.J. (1987). Symbiotic and competitive properties of motility mutants of *Rhizobimn trifoii* TAl. Arch. Microbiol., 148: 34-39.

Mendoza-Suárez, M.A., Geddes, B.A., Sánchez-Cañizares, C., Ramírez-González, R.H., Kirchhelle, C., Jorrin, B., Poole, P.S. (2020). Optimizing *Rhizobium*-legume symbioses by simultaneous measurement of rhizobial competitiveness and N_2 fixation in nodules. Proc. Natl Acad. Sci. USA, 117: 9822-9831.

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., Kondorosi, E. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*–legume symbiosis. Proc. Natl Acad. Sci. USA 103: 5230-5235.

Moawad, M., Schmidt, E.L. (1987). Occurrence and nature of mixed infections in nodules of field-grown soybeans (*Glycine max*). Biol. Fert. Soils, 5: 112-114.

Mongiardini, E.J., Pérez-Giménez, J., Althabegoiti, M.J., Covelli, J., Quelas, J.I., López-García, S.L., Lodeiro, A.R. (2009). Overproduction of the rhizobial adhesin RapA1 increases competitiveness for nodulation. Soil Biol. Biochem., 41: 2017-2020.

Muller, K.E., Denison, R.F. (2018). Resource acquisition and allocation traits in symbiotic rhizobia with implications for life-history outside of legume hosts. R. Soc. Open Sci., 5: 181-124.

Murphy, P.J., Wexler, W., Grzemski, W., Rao, J.P., Gordon, D. (1995). Rhizopines – their role in symbiosis and competition. Soil Biol. Biochem., 27: 525-529.

O'Brian, M.R., Kirshbom, P.M., Maier, R.J. (1987). Bacterial heme synthesis is required for expression of the leghemoglobin holoprotein but not the apoprotein in soybean root nodules. Proc. Natl. Acad. Sci. USA, 84: 8390-8393.

Okazaki, S., Yuhashi, K., Minamisawa, K. (2003). Quantitative and time-course evaluation of nodulation competitiveness of rhizobitoxine-producing *Bradyrhizobium elkanii*. FEMS Microbiol. Ecol., 45: 155-160.

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

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

Olivieri, I., Frank, S.A. (1994). The evolution of altruism in rhizobium: altruism in the rhizosphere. J. Hered., 85: 46-47.

Oono, R., Anderson, C.G., Denison, R.F. (2011). Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. Proc. Biol. Sci., 278: 2698-2703.

Oono, R., Denison, R.F., Kiers, E.T. (2009). Tansley review: Controlling the reproductive fate of rhizobia: how universal are legume sanctions? New Phytol., 183: 967-979.

Ott, T. van Dongen, J.T., Gunther, C., Krusell, L., Desbrosses, G., Vigeolas, H., Bock, V., Czechowski, T., Geigenberger, P., Udvardi, M.K. (2005). Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. Curr. Biol., 15: 531-535.

Parco, S.Z., Dilworth, M.J., Glenn, A.R. (1994). Motility and the distribution of introduced root nodule bacteria on the root system of legumes. Soil Biol. Biochem., 26: 297-300.

Patriarca, E.J., Tate, R., Iaccarino, M. (2002). Key role of bacterial NH_4^+ metabolism in *Rhizobium*-plant symbiosis. Microbiol. Mol. Biol. Rev., 66: 203-22.

Pérez-Giménez, J., Althabegoiti, M.J., Covelli, J., Mongiardini, E.J., Quelas, J.I., López-García S.L., Lodeiro, A.R. (2009). Soybean lectin enhances biofilm formation by *Bradyrhizobium japonicum* in the absence of plants. Int. J. Microbiol., 2009: 1-9.

Pérez-Giménez, J., Quelas, J. I., Lodeiro, A. R. (2011). "Competition for nodulation" in Physiology and Biochemistry, (ed.) El-Shemy, H. A. (Rijeka: InTech Open Access Publisher), 139-166.

Pérez-Montaño, F., Jiménez-Guerrero, I., Acosta-Jurado, S., Navarro-Gómez, P, Ollero, F.J., Ruiz-Sainz, J.E., López-Baena, F.J., Vinardell, J.M. (2016). A transcriptomic analysis of the effect of genistein on *Sinorhizobium fredii* HH103 reveals novel rhizobial genes putatively involved in symbiosis. Sci. Rep., 6: 31592.

Peters, J.W., Szilagyi, R.K. (2006). Exploring new frontiers of nitrogenase structure and mechanism. Curr. Opin. Chem. Biol., 10: 101-8.

Pierre, O., Engler, G., Hopkins, J., Brau, F., Boncompagni, E., Herouart, D. (2013). Peribacteroid space acidification: a marker of mature bacteroid functioning in *Medicago truncatula* nodules. Plant Cell Environ., 36: 2059-2070.

Pini, F., East, A.K.; Appia-Ayme, C.; Tomek, J.; Karunakaran, R.; Mendoza-Suárez, M.; Edwards, A.; Terpolilli, J.J.; Roworth, J.; Downie, J., Poole, P.S. (2017). Bacterial biosensors for *in vivo* spatiotemporal mapping of root secretion. Plant Physiol., 174: 1289-1306.

Pobigaylo, N., Szymczak, S., Nattkemper, T.W., Becker, A. (2008). Identification of genes relevant to symbiosis and competitiveness in *Sinorhizobium meliloti* using signature-tagged mutants. Mol. Plant-Microbe Interact., 21: 219-231.

Poole, P.S., Allaway, D. (2000). Carbon and nitrogen metabolism in *Rhizobium*. Adv. Microb. Physiol. 43: 117–163

Poole, P.S., Ramachandran, V., Terpolilli, J. (2018). Rhizobia: from saprophytes to endosymbionts. Nat. Rev. Microbiol., 16: 291-303.

Porter, S.S., Chang, P.L., Conow, C.A., Dunham, J.P., Friesen, M.L. (2017). Association mapping reveals novel serpentine adaptation gene clusters in a population of symbiotic *Mesorhizobium*. ISME J., 11: 248-262.

Potts, S.G., Biesmeijer, J.C., Kremen, C., Neumann, P., Schweiger, O., Kunin, W.E. (2010). Global pollinator declines: trends, impacts and drivers. Trends Ecol. Evol., 25: 345-353.

Prell, J., Bourdes, A., Kumar, S., Lodwig, E., Hosie, A., Kinghorn, S., White, J., Poole, P.S (2010). Role of symbiotic auxotrophy in the *Rhizobium*-legume symbioses. PLoS One, 5: e13933.

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

Pueppke, S.G., Broughton, W.J. (1999). *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. Mol. Plant Microbe. Interact., 12: 293-318.

Puppo, A., Groten, K., Bastian, F., Carzaniga, R., Soussi, M., Lucas, M.M., de Felipe, M.R., Harrison, J., Vanacker, H., Foyer, C.H. (2005). Legume nodule senescence: roles for redox and hormone signalling in the orchestration of the natural aging process. New Phytol., 165: 683-701.

Quelas, J.I., López-García, S.L., Casabuono, A., Althabegoiti, M.J., Mongiardini, E.J., Pérez-Giménez, J., Couto, A., Lodeiro, A.R. (2006). Effects of N-starvation and C-source on exopolysaccharide production and composition in *Bradyrhizobium japonicum*. Arch. Microbiol., 186: 119-128.

Quelas, J.I., Mongiardini, E.J., Casabuono, A., López-García, S.L., Althabegoiti, M.J., Covelli, J.M., Pérez-Giménez, J., Couto, A., Lodeiro, A.R. (2010). Lack of galactose or galacturonic acid in *Bradyrhizobium japonicum* USDA 110 exopolysaccharide leads to different symbiotic responses in soybean. Mol. Plant Microbe Interact., 23: 1592-1604.

Raetz, C.R.H., Whitfield, C. (2002). Lipopolysaccharide endotoxins. Annu. Rev. Biochem., 71: 635-700.

Rajwar, A., Sahgal, M., Johri, B.N. (2013). "Legume–rhizobia symbiosis and interactions in agroecosystems" in Plant Microbe Symbiosis: Fundamentals and Advances, (ed.) Arora, N., (Springer: New Delhi, India), 233-265.

Ramachandran, V.K., East, A.K., Karunakaran, R., Downie, J.A., Poole, P.S. (2011). Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biol., 12: 106-109.

Ratcliff, W.C., Denison, R.F. (2009). Rhizobitoxine producers gain more poly-3-hydroxybutyrate in symbiosis than do competing rhizobia, but reduce plant growth. ISME J., 3: 870-872.

Ratcliff, W.C., Kadam, S., Denison, R.F. (2008). Poly-3-hydroxybutyrate (PHB) supports survival and reproduction in starving rhizobia. Fems Microbiol. Ecol., 65: 391-399.

Raymond, J., Siefert, J.L., Staples, C.R., Blankenship, R.E. (2004). The natural history of nitrogen fixation. Mol. Biol. Evol., 21: 541-554.

Regus, J. U., Gano, K. A., Hollowell, A.C., Sofish, V., Sachs, J.L. (2015). *Lotus* hosts delimit the mutualism-parasitism continuum of *Bradyrhizobium*. J. Evol. Biol., 28: 447-456.

Regus, J., Gano, K.A., Hollowell, A.C., Sachs, J.L. (2014). Efficiency of partner choice and sanctions in *Lotus* is not altered by nitrogen fertilization. Proc. R. Soc. B., 281: 20132587-20132587.

Regus, J.U., Quides, K.J., O'Neill, M.R., Suzuki, R., Savory, E.A., Chang, J.H., J. Sachs, J.L. (2017). Cell autonomous sanctions in legumes target ineffective rhizobia in nodules with mixed infections. Am. J. Bot., 104: 1299-1312.

Relic, B., Talmont, F., Kopcinska, J., Golinowsky, W., Prome´, J.C., Broughton, W.J. (1993). Biological activity of *Rhizobium* sp NGR234 nod-factors on *Macroptilium atropurpureum*. Mol. Plant. Microbe Interact., 6: 764-774.

Remigi, P., Capela, D., Clerissi, C., Tasse, L.N., Torchet, R., Bouchez, O., Batut, J., Cruveiller, S., Rocha, E.P.C., Masson-Boivin, C. (2014). Transient hypermutagenesis accelerates the evolution of legume endosymbionts following horizontal gene transfer. PLoS Biol. 12: e1001942.

Remigi, P., Zhu, J., Young, J.P., Masson-Boivin, C. (2016). Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. Trends Microbiol., 24: 63-75.

Robleto E.A., Kmiecik K., Oplinger E.S., Nienhaus J., Triplett E.W. (1998). Trifolitoxin production increases nodulation competitiveness of *Rhizobium etli* CE3 under agricultural conditions. Appl. Environ. Microbiol., 64: 2630-2633.

Roy, S., Liu, W., Nandety, R.S., Crook, A., Mysore, K.S., Pislariu, C.I., Frugoli, J., Dickstein, R., Udvardi, M.K. (2020). Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. Plant Cell, 32: 15-41.

Sachs, J.L., Ehinger, M., Simms, E. (2010a). Origins of cheating and loss of symbiosis in wild *Bradyrhizobium*. J. Evol. Biol., 23: 1075-1089.

Sachs, J.L., Mueller, U.G., Wilcox, T.P., Bull, J.J. (2004). The evolution of cooperation. Q. Rev. Biol., 79: 135-160.

Sachs, J.L., Quides, K.W., Wendlandt, C.E. (2018). Legumes versus rhizobia: a model for ongoing conflict in symbiosis. New Phytol., 219: 1199-1206.

Sachs, J.L., Russell, J.E., Lii, Y.E., Black, K.C., Lopez, G., Patil, A.S. (2010b). Host control over infection and proliferation of a cheater symbiont. J. Evol. Biol., 23: 1919-1927.

Sachs, J.L., Simms, E.L. (2006). Pathways to mutualism breakdown. Trends Ecol. Evol., 21: 585-592.

Sadowsky, M.J., Graham, P.H. (1998). "Soil biology of the *Rhizobiaceae",* in Molecular biology of model plant-associated bacteria, (eds.) Spaink H.P., Kondorosi A., Hooykaas P.J.J. (Kluwer Academic Publishers), 155-172.

Scheidle, H., Gross, A., Niehaus, K. (2005). The lipid A substructure of the *Sinorhizobium meliloti* lipopolysaccharides is sufficient to suppress the oxidative burst in host plants. New Phytol., 165: 559-65.

Schumpp, O., Deakin, W. J. (2010). How inefficient rhizobia prolong their existence within nodules. Trends Plant Sci., 15: 189-195.

Segovia, L., Pinero, D., Palacios, R., Martinez-Romero, E. (1991). Genetic structure of a soil population of nonsymbiotic *Rhizobium leguminosarum*. Appl. Environ. Microbiol., 57: 426-433.

Sheppard, S.K., Didelot, X., Meric, G., Torralbo, A., Jolley, K.A., Kelly, D.J., Bentley, S.D., Maiden, M.C., Parkhill, J., Falush, D. (2013). Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in *Campylobacter*. Proc. Natl. Acad. Sci., 110: 11923-11927.

Simms, E.L., Bever, J.D. (1998). Evolutionary dynamics of rhizopine within spatially structured rhizobium populations. Proc. R. Soc. Lond. Ser. B., 265: 1713-1719.

Simms, E.L., Taylor, D.L., Povich, J., Shefferson, R.P., Sachs, J.L., (2006). An empirical test of partner choice mechanisms in a wild legume-rhizobium interaction. Proc. R. Soc. Lond. Ser. B., 273: 77-81.

Singleton, P.W., Stockinger, K.R. (1983). Compensation against ineffective nodulation in soybean. Crop Sci., 23, 69-72.

Smil, V. (2001). Enriching the Earth: Fritz Haber, Carl Bosch and the Transformation of World Food Production, The MIT Press: Cambridge, MA.

Smith, J.M. (1989). Generating novelty by symbiosis. Nature., 341: 284-285.

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

Sullivan, J.T., Eardly, B.D., van Berkum, P. and Ronson, C.W. (1996). Four unnamed species of nonsymbiotic rhizobia isolated from the rhizosphere of *Lotus corniculatus*. Appl. Environ. Microbiol., 62: 2818-2825.

Tellstrom, V., Usadel, B., Thimm, O., Stitt, M., Kuster, H., Niehaus, K. (2007). The lipopolysaccharide of *Sinorhizobium meliloti* suppresses defense-associated gene expression in cell cultures of the host plant *Medicago truncatula*. Plant Physiol., 143: 825-37.

Terpolilli, J.J., Masakapalli, S.K., Karunakaran, R., Webb, I., Green, R., Watmough, N.J., Kruger, N.J., Ratcliffe, R.G., Poole, P.S. (2016). Lipogenesis and redox balance in nitrogen-fixing pea bacteroids. J. Bacteriol, 198: 2864-2875

Thies, J.E., Singleton, P.W., Bohlool, B.B. (1991b). Modelling symbiotic performance of introduced rhizobia in the field by use of indices of resident population size and nitrogen status of the soil. Appl. Environ. Microbiol., 57: 29-37.

Thrall, P., Slattery, J., Broadhurst, L., Bickford, S. (2007). Geographic patterns of symbiont abundance and adaptation in native Australian *Acacia*–rhizobia interactions. J. Ecol., 95: 1110-1122.

Toro, N. (1996). Nodulation competitiveness in the *Rhizobium* legume symbiosis. World J. Microbiol. Biotech., 12: 157-162.

Trainer, M.A., Charles, T.C. (2006). The role of PHB metabolism in the symbiosis of rhizobia with legumes. Appl. Microbiol. Biotechnol., 71(4): 377-386.

Triplett, E.W., Sadowsky, M. (1992). Genetics of competition for nodulation of legumes. Annu. Rev. Microbiol., 46: 399-428.

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

Van Berkum, P., Elia, P., Song, Q., Eardly, B.D. (2012). Development and application of a multilocus sequence analysis method for the identification of genotypes within genus *Bradyrhizobium* and for establishing nodule occupancy of soybean (*Glycine max* L. Merr). Mol. Plant Microbe Interact., 25: 321-330.

van de Velde, W., Guerra, J.C., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., Mergaert, P., Kondorosi, E., Holsters, M., Goormachtig, S. (2006). Aging in legume symbiosis. A molecular view on nodule senescence in *Medicago truncatula*. Plant Physiol., 141: 711-720.

van Dillewijn, P., Soto, M.J., Villadas, P.J., and Toro, N. (2001). Construction and environmental release of a *Sinorhizobium meliloti* strain genetically modified to be more competitive for alfalfa nodulation. Appl. Environ. Microbiol., 67: 3860-3865.

Vance, C.P. (2000). "Amide biosynthesis in root nodules of temperate legumes" in Prokaryotic nitrogen fixation: a model system for analysis of a biological process, (ed.) Triplett, E.W. (Wymondham, UK: Horizon Scientific Press), 589-608.

Vanderlinde, E.M. Muszynski, A., Harrison, J.J., Koval, S.F., Foreman, D.L., Ceri, H., Kannenberg, E.L., Carlson, R.W., Yost, C.K. (2009). *Rhizobium leguminosarum* biovar *viciae* 3841, deficient in 27-hydroxyoctacosanoate-modified lipopolysaccharide, is impaired in desiccation tolerance, biofilm formation and motility. Microbiol., 155: 3055-3069.

Vasse, J., de Billy, F., Camut, S., Truchet, G. (1990). Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. J. Bacteriol., 172: 4295-306.

Velázquez, E., García-Fraile, P., Ramírez-Bahena, M.H., Rivas, R., Martínez-Molina, E. (2010). Bacteria involved in nitrogen-fixing legume symbiosis: Current taxonomic perspective, in Microbes for legume improvements, (eds) Khan M.S., Musarrat J., Zaidi A., Springer, 1-25.

Wang, D., Wang, Y.C., Wu, L.J., Liu, J.X., Zhang, P., Jiao, J., Yan, H., Liu, T., Tian, C.F., Chen, W.X. (2016). Construction and pilot screening of a signature-tagged mutant library of *Sinorhizobium fredii*. Arch. Microbiol., 198: 91-99.

Wang, D., Yang, S., Tang, F., Zhu, H (2012). Symbiosis specificity in the legume-rhizobial mutualism. Cell Microbiol., 14: 334-342.

West, S.A., Griffin, A.S., Gardner, A. (2007). Evolutionary explanations for cooperation. Curr. Biol., 17: R661-R672

West, S.A., Kiers, E.T., Simms, E.L., Denison, R.F. (2002). Sanctions and mutualism stability: why do rhizobia fix nitrogen? Proc. R. Soc. Lond. B., 269: 685-694.

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

Wheatley, R.M., Poole, P.S. (2018). Mechanisms of bacterial attachment to roots. FEMS Microbiol. Rev., 42: 448-461.

Wilson, G.W.T., Rice, C.W., Rillig, M.C., Springer, A., Hartnett, D.C. (2009). Soil aggregation and carbon sequestration are tightly correlated with the abundance of arbuscular mycorrhizal fungi: results from long-term field experiments. Ecol. Lett., 12: 452-461.

Xu, J., Webb, I., Poole, P.S., Huang, W.E. (2017). Label-free discrimination of Rhizobial bacteroids and mutants by single-cell Raman microspectroscopy. Anal. Chem., 89 (12): 6336-6340.

Yates, R.J., Howieson, J.G., Reeve, W.G., O'Hara, G.W. (2011). A re-appraisal of the biology and terminology describing rhizobial strain success in nodule occupancy of legumes in agriculture. Plant Soil, 348: 255-267.

Young, J.P.W. (2016). Bacteria are smartphones and mobile genes are apps. Trends Microbiol., 24: 931-932.

Chapter 2

Aim of the thesis

A major task, which humanity will challenge in the forthcoming future, is to sustain the increased demand of food without altering the integrity of the Earth's environmental systems. Nowadays, the model of agricultural production is yet depending on the use of nitrogen fertilizers. Their massive use has introduced in the ecosystem a large amount of reactive nitrogen, causing severe environmental problems. However, in the last years the employment of eco-sustainable agronomic practices is progressively increased. Moreover, the ecological and economical importance of the legume-rhizobia symbiosis has encouraged the production of commercial rhizobia inoculants, as inexpensive alternatives to industrial nitrogen fertilizers that can improve legume yields. Indeed, the biological N_2-fixation (BNF) can be maximized in agricultural systems by the selection of high nitrogen fixation efficient rhizobia strains to improve the productivity of leguminous plants. However, to date, the effectiveness of most of the employed rhizobial inoculants can be limited by the poor ability of bio-inoculants to adapt to harshly soil conditions and by the 'competition problem' that arise in agricultural settings. In this perspective, to ameliorate the selection process of strains for further production of rhizobial inoculants, the main goal of this thesis is to discover which are the *Sinorhizobium meliloti* genomic features would increase its competitiveness in the rhizosphere among other rhizobia. By this purpose, in this thesis an innovative strategy, which combine newly genome-wide association (GWAS) approach with large phenotypic screening, is reported. This strategy could be in the future successfully applied for extensive evaluations of rhizobila inoculants.

The selection and a deep characterization of rhizobial strains is also crucial to develop efficient legume-growing systems to be employed as eco-sustainable agronomic practice. Thus, using traditional microbiological approaches will be applied to identify rhizobial strains able to establish an efficient symbiosis and thus increase the productivity of two *Pisum sativum* emerging breeding lines, and its related phenotypic features for improved adaptation to different soil conditions.

Chapter 3

Competitiveness prediction for nodule colonization in *Sinorhizobium meliloti* through combined *in vitro* tagged strain characterization and genome-wide association analysis

A. Bellabarba [1,2], G. Bacci[3], F. Decorosi[1,2], E. Aun[4], E. Azzarello[1], M. Remm[4], L. Giovannetti [1,2], C. Viti [1,2*], A. Mengoni[3*] and F. Pini[5]

[1]Department of Agronomy, Food, Environmental and Forestry (DAGRI), University of Florence, Via della Lastruccia 10, I-50019, Sesto Fiorentino (Italy)

[2]Genexpress Laboratory, Department of Agronomy, Food, Environmental and Forestry (DAGRI), University of Florence, Via della Lastruccia 14, I-50019, Sesto Fiorentino, Italy

[3]Department of Biology, University of Florence, Via Madonna del Piano 6, I-50019 Sesto Fiorentino, Italy

[4] Department of Bioinformatics, Institute of Molecular and Cell Biology, University of Tartu, Riia 23, 51010 Tartu, Estonia

[5] Department of Biology, University of Bari Aldo Moro, via Orabona 4, I-70124 Bari, Italy

Manuscript submitted on mSystems and on bioRxiv (doi:https://doi.org/10.1101/2020.09.15.298034)
Personal contribution: conceived and designed the experiments, performed the experiments, analyzed the data and wrote the paper

Competitiveness prediction for nodule colonization in *Sinorhizobium meliloti* through combined *in vitro* tagged strain characterization and genome-wide association analysis

A. Bellabarba[1,2], G. Bacci[3], F. Decorosi[1,2], E. Aun[4], E. Azzarello[1], M. Remm[4], L. Giovannetti [1,2], C. Viti [1,2*], A. Mengoni[3*] and F. Pini[5]

[1]Department of Agronomy, Food, Environmental and Forestry (DAGRI), University of Florence, Via della Lastruccia 10, I-50019, Sesto Fiorentino (Italy)

[2]Genexpress Laboratory, Department of Agronomy, Food, Environmental and Forestry (DAGRI), University of Florence, Via della Lastruccia 14, I-50019, Sesto Fiorentino, Italy

[3]Department of Biology, University of Florence, Via Madonna del Piano 6, I-50019 Sesto Fiorentino, Italy

[4] Department of Bioinformatics, Institute of Molecular and Cell Biology, University of Tartu, Riia 23, 51010 Tartu, Estonia

[5] Department of Biology, University of Bari Aldo Moro, via Orabona 4, I-70124 Bari, Italy

* **Corresponding authors**

Alessio Mengoni, Department of Biology, University of Florence, via Madonna del Piano 6, 50019, Sesto Fiorentino, Italy. Tel.: +39 055 4574738. E-mail: alessio.mengoni@unifi.it

Carlo Viti, Department of Agronomy, Food, Environmental and Forestry (DAGRI), University of Florence, Via della Lastruccia 10, I-50019, Sesto Fiorentino, (Italy). Tel.: +39 055 4573248. E-mail: carlo.viti@unifi.it

Date submitted 15/09/2020

Running title: Genome-wide association study for efficient prediction of competitiveness in *Sinorhizobium meliloti*

ABSTRACT

Associations between leguminous plants and symbiotic nitrogen fixing bacteria (rhizobia) are ac lassical example of mutualism between a eukaryotic host and a specific group of prokaryotic microbes. Though being in part species-specific, different strains may colonize the same plant symbiotic structure (nodule). It is known that some rhizobial strains are better competitor than others, but detailed analyses aimed to predict from the rhizobial genome its competitive abilities are still scarce. Here we performed a bacterial genome wide association (GWAS) analysis to define the genomic determinants related to the competitive capabilities in the model rhizobial species *Sinorhizobium meliloti.* Thirteen tester strains were GFP-tagged and assayed against three reference competitor strains RFP-tagged (Rm1021, AK83 and BL225C) in a *Medicago sativa* nodule occupancy test. Competition data in combination with strains genomic sequences were used to build-up a model for GWAS based on k-mers. The model was then trained and applied for competition capabilities prediction. The model was able to well predict the competition abilities against two partners, BL225C, Rm1021 with coefficient of determination of 0.96 and 0.84, respectively. Four strains showing the highest competition phenotypes (> 60% single strain nodule occupancy; GR4, KH35c, KH46 and SM11) *versus* BL225C were used to identify k-mers associated with competition. The k-mers with highest scores mapped on the symbiosis-related megaplasmid pSymA and on genes coding for transporters, proteins involved in the biosynthesis of cofactors and proteins related to metabolism (i.e. glycerol, fatty acids) suggesting that competition abilities reside in multiple genetic determinants comprising several cellular components.

IMPORTANCE Decoding the competitive pattern that occurs in the rhizosphere is challenging in the study of bacterial social interaction strategies. To date, single-gene approach has been mainly used to uncover the bases of nodulation, but there is still a gap about the main features that *a priori* turn out rhizobial strains able to outcompete indigenous rhizobia. Therefore, tracking down which traits make different rhizobial strains able to win the competition for plant infection over other indigenous rhizobia will allow ameliorating strain selection and consequently plant yield in sustainable agricultural production systems.

We have proven that a k-mer based GWAS approach can effectively predict the competition abilities of a panel of strains, which were analyzed for their plant tissue occupancy by using double fluorescent labeling. The reported strategy could be used for detailed studies on the genomic aspects of the evolution of bacterial symbiosis and for an extensive evaluation of rhizobial inoculants.

INTRODUCTION

The nitrogen-fixing symbiotic interaction is a classic example of a mutualistic association (1). The interaction between rhizobia and host plants (mostly *Fabaceae*, but also with the genus *Parasponia*) starts with an exchange of molecular signals: flavonoids released from the plant roots and Nod factors produced by rhizobia (2). Nod factors induce a molecular pathway on plant root cells, which ultimately leads to rhizobia entry in the tissue and intracellular colonization. Intracellular rhizobia differentiate into bacteroids, which express the nitrogenase, the enzyme responsible for the fixation of atmospheric di-nitrogen to ammonia (3, 4). On the plant root, a new structure, called nodule, forms, where rhizobia can proliferate and differentiate into bacteroids (5, 6). Indeed, in a single nodule (a mass of a few hundreds of milligrams), up to 10^6 bacterial cells can be recovered, while in the soil, free-living rhizobia are usually not more than 10^3-10^4/g of soil (7).

As in a trade framework, the benefit for the rhizobium is a protected environment where it can proliferate (under control) and receive carbon and energy supply from the plant, while the reward for the plant is the availability of fixed nitrogen (8, 9). Rhizobia transmission is horizontal among plants, and plants may be colonized by poorly effective (low nitrogen fixing) strains. However, host plants could control the spread of inefficient strains by sanctioning root nodules, hence limiting the growth of underperforming strains (10–12). Moreover, the presence of multiple strains within a nodule (mixed nodule) may occur (7, 13). Under these circumstances, inefficient strains can behave as cheaters decreasing the overall nitrogen-fixing performance (7, 8). Consequently, understanding the competition between strains has a great importance either to study the evolution of symbiosis (14, 15) and to predict how much efficient would be a rhizobial inoculant for agricultural applications (16). The genetic bases of competitiveness among rhizobial strains are still elusive and, most of the studies, addressed symbiotic genetic determinants from experiments involving mutants of a few different natural strains (7).

The link between a phenotype and its genetic basis, hence predicting phenotypes from the sole genomic information is one of the great challenges of biology (17). Genome-Wide Association Studies (GWAS) are commonly used for identifying the putative functional role of set of allelic variation in groups of individuals. In bacteria, GWAS have been applied to several species for predicting complex (i.e. multigenic) phenotypes, as antibiotic tolerance and host interaction (see for examples (18–20). However, most studies were related to phenotypes under strong selective pressure, while it is still challenging to determine the genetic basis of phenotypes under mild selection (21). The identification of the genetic determinants in host-bacteria interaction is essential, and GWASs on plant holobionts (the ensemble of the plant and the other organisms living in or around) have been proposed (21), aiming to provide the basis for future breeding programs, which includes among the plant trait, the recruitment of the "good" microbiome.

Recent studies have reported the feasibility of experimental setups combining symbiotic assays with genome sequencing approaches and GWAS in rhizobia to define the genetic determinants of symbiotic performances (24–27). In the model rhizobium *Sinorhizobium meliloti*, association analysis has been used to explore the genetic basis of various phenotypic traits including antibiotic resistance, symbiotic and metabolic traits (24). A select and re-sequence approach has been successfully applied to measure the fitness of a set of 101 *S. meliloti* strains against two genotypes of the host plant *Medicago truncatula* (25). However, the predictive value of single rhizobial genotypes (i.e. genomes) toward the expected fitness in terms of competitive capabilities to establish successful symbiosis is still unclear. Indeed, to date, though many genetic details of the symbiotic interaction are known for single strain colonization (28), there is still a gap about which rhizobial features increase the chances to win the competition for plant infection outcompeting others indigenous rhizobia. Therefore, unearthing these genetic determinants can promote studies on the genomic aspects of the evolution of bacterial symbiosis and can have direct practical application in the screening and amelioration of rhizobial inoculants to be used in sustainable agricultural production systems.

Here, we aimed to address the possibility to predict, on the basis of the sole genome sequence, the competitive capabilities of rhizobial strains. We used as model rhizobial species, *Ensifer* (syn. *Sinorhizobium*) *meliloti* for which many molecular genetics data and tools are available (29), a good number of strains has been sequenced and preliminary data on symbiotic performances and competition are present (7, 18, 19, 25, 30). We settled a series of experiments where pairs of fluorescently labelled *S. meliloti* strains were used to infect alfalfa (*Medicago sativa*) plants. The measure of the competition phenotypes of each strain was then used with

genomic sequences of the same strain to perform genome-wide association analysis and construct a predictive model of competition based on genomic features (k-mers) occurrence.

RESULTS

Construction of fluorescently tagged *S. meliloti* strains. To set up *in vitro* tests for measuring competition capabilities, a panel of 16 *S. meliloti* strains was selected. Three well characterized strains for competition capabilities (*S. meliloti* BL225C, AK83 and Rm1021) were chosen (7) and used as reference competitors against 13 *S. meliloti* strains (tester strains) whose genome sequences were available (Table S1). Phylogenetic relationships among the 13 *S. meliloti* strains were evaluated (Fig. S1A) and their pangenome was analyzed (Fig. S1B-D). The pangenome is composed of 15419 genes: 4278 are shared by all strains (core genome) while 6622 genes are strain specific (Fig. S1D). For all the above-mentioned 13 tester strains, GFP (green fluorescent protein) derivatives were constructed by cloning the pHC60 plasmid, which constitutively expresses the GFP. *S. meliloti* Rm1021, BL225C and AK83 strains were tagged with RFP (red fluorescent protein) by using the pBHR mRFP plasmid. Preliminary single inoculation assays were performed showing that all strains were able to form nodules on the root of *M. sativa* plants (Fig. S2). For all but two strains (M270 and T073) nitrogenase activity inside nodules was detected (Fig. S2D), in agreement with previous results that showed low nitrogen fixation abilities in symbiotic interaction with *M. truncatula* (31). Differences in nodulation, plant growth and nitrogenase activity among strains were observed. *S. meliloti* AK58, RU11/001, SM11, USDA1157, GR4, and CCMM B554 strains gave the highest values of nitrogenase activity and plant growth promotion (Fig. S2).

Competition capabilities for colonization of nodules differ in relation to competitor counterpart. Tagged *S. meliloti* strains were used in a set of competition experiments: each with a GFP-tagged strain (13 strains in total) versus an RFP-tagged competitor strain (*S. meliloti* BL225C, AK83 or Rm1021) (13 x 3 total of 39). Large variability in nodule colonization was observed among and within the three sets of competitions experiments (vs Rm1021, vs AK83 and vs BL225C). The three competitions test showed differences in the number of total nodules produced per plant ($p < 0.001$, Fig. S3A) and, in the competition experiments with AK83 the highest number of nodules was observed (Fig. S3A).

Competition capabilities were evaluated as single nodule occupancy (nodules colonized by a single strain) of the tested strain in respect to a reference strain; good competitors were

characterized by a single nodule occupancy higher than 60%, medium competitor between 20 and 60% and weak competitors below 20%. In the competition with Rm1021, most of the strains outperformed (Fig. 1A). This competition test was characterized by a high average value of nodule occupancy of the strains tested equal to 65.12%. Most of the strains showed single nodule occupancy higher than 60%, with the highest value (93.4%) observed for GR4, three strains (HM006, Rm41 and M270) exhibited medium competition capabilities and two strains (T073 and USDA1157) lowly performed (Fig. 1A and Table S2). Among the three strains with medium competition capabilities, HM006 and Rm41 displayed values close to high performing strains (58.6% and 50.5%, respectively) while M270 nodule occupancy was 37%. Medium-low performance of strains M270 and T073 (T073 competition was characterized by mixed nodules or with strain Rm1021 only) may be related to nodule sanctioning (plant limiting nutrient to inefficient nodules) as T073 and M270 were unable to fix nitrogen (Fig. S2D). Concerning USDA1157 strain, which showed good nitrogen-fixing abilities, we might hypothesize that the lower value of nodule occupancy could be related to direct strain-by-strain competitive interaction, more than to plant sanctions.

In the competition experiments with AK83, conversely to the pattern highlighted when Rm1021 was used as a competitor, a general decrease of nodule occupancy of the 13 strains tested was observed (Fig. 1B), resulting in a lower average value of nodule occupancy (21.99%; Fig. 1B and Table S2). Except for GR4, that showed the highest percentage of occupancy equal to 63.9%, all strains displayed weak-medium competitive capabilities (nodule occupancy lower than 60%; Fig. 1B and Table S2). The lowest value of nodule occupancy equal to 1.7% and 1.8% were detected for Rm41 and T073, respectively.

An intermediate scenario between those above described was detected for the competition with BL225C (Fig. 1C and Table S2), where the average value of nodule occupancy of the strains tested was equal to 39.45%. The most competitive strains were GR4, KH35c, KH46, and SM11, showing a nodule occupancy ranging from 63.4 % to 68.9%. The lowest percentage of nodule occupancy equal to 0.4% was detected for T073.

In both the competition with AK83 and BL225C, a higher abundance of mixed nodules (nodules infected by both *S. meliloti* strains) was observed compared to the competition with Rm1021 (Fig. S3B).

Modeling competition pattern from genome sequences. Short DNA oligomers with constant length k, termed k-mers, allow to capture a large set of genetic variants in a population, including SNPs, insertions/deletions (INDELs) (20, 32). To pinpoint genetic determinants that might be responsible for the competition capabilities variation among *S. meliloti* strains, we identified competitive phenotype-specific k-mers and trained a k-mer-based statistical model for predicting phenotypes of interest.

A matrix based on the competition phenotype (Table S2) and the genome sequences of the strains tested, was used to identify the most strongly associated-phenotype k-mers (p-value < 0.05) and linear regression models were built for each set of competition experiments. To measure the accuracy of predictive models, the strains were randomly split into a training test and a test set (32).

The effectiveness of linear regression models in describing the three observed competition patterns varied among the dataset. The prediction was greatly accurate for the competition phenotype versus BL225C (Fig. 2C), as indicated by the coefficient of determination R^2 equal to 0.98 and 0.96 in training and test set respectively (Table 1), and by the degree of similarity between actual and predicted phenotype (Fig. 2), The number of total k-mers identified in tested strains as significantly associated with competition against BL225C was smaller in comparison with other datasets (Table 1). The wider p-value range for significantly associated k-mers was founded in the analysis of competition against BL225C.

In terms of model-evaluation metrics, the linear regression model was less reliable in defining the competition against Rm1021 for both training and test sets respect to the competition against BL225C (Table 1). In particular, the predictive capacity of the model was inaccurate for those strains that displayed low or medium-low competition capabilities (*S. meliloti* T073, USDA1157, M270, Rm41 and HM006) in a competition where most of the strains showed medium-high capacities (Fig. 2). Furthermore, due to the greater extent of positive performances showed by most of the strains tested against Rm1021, a higher amount of total k-mers related to this competition phenotype was obtained (Table 1). Concerning the level of accuracy of the predictive model for competition against AK83, the results appeared discordant (Table 1). Indeed, the model well explained the competition phenomenon only for those strains that belong to the training set ($R^2 = 0.99$), while it was unreliable for the test set ($R^2 = 0.11$). Specifically, AK58 and GR4 performances were underestimated (Fig. 2B). An intermediate amount of "common" k-mers significantly associated with competition phenotype against AK83 was achieved (Table 1).

The modeling analyses were also trained, creating three different random combinations of training and test set. The coefficient of determination R^2 of the linear regression models of the Rm1021 and AK83 datasets significantly changed compared to the previous analyses, indicating that the models were strongly influenced by the random split of samples into training and test set (Table S3). On the contrary, the coefficients of determination R^2 of BL225C dataset were less affected by the different combinations of training/test set, indicating that this competition phenotype was well represented by the model for both training and test sets (Table S3). Overall, these results suggest that the linear regression model can well predict the most contrasting competition phenotype in BL225C.

Putative genetic determinants associated with good competition capabilities against *S. meliloti* BL225C. The linear regression model for the competition against BL225C was the most reliable; therefore, we focused on the set of k-mers related to this dataset. We produced a list of k-mers significantly associated with competition phenotypes covering putative genetic variants. K-mers (p-value < 0.05) were mapped on the annotated genomes of 13 tested strains to retrieve the genetic determinants associated with the phenotype of interest. Among the top k-mers (p-value threshold < 0.0001, Table S4), 51 k-mers (with a p-value equal to 1.31 e^{-04}) mapped in genomes of the four strains that showed single nodule occupancy higher than 60% (*S. meliloti* GR4 SM11 KH35c and KH46) in the competition test against BL225C (Table S4, highlighted in bold). Unlike other top k-mers, the coefficients in the linear regression model of these 51 k-mers have all positive values, indicating that there is a positive correlation between the presence of k-mer found in the sample and the phenotypic value attestation. Consequently, based on the above evidence, we focused on these 51 k-mers to define the genetic determinants putatively associated with good competition capabilities against BL225C. These k-mers tagged 103 predicted protein-coding sequences (CDS) (one k-mer can tag multiple genes, Table S5). Among the 103 CDSs, a set of orthologous genes was identified (Table S4). These orthologous genes hits were mostly tracked in *S. meliloti* GR4, KH35c and KH46 genomes (Fig. S4B) and were predominantly located on the symbiosis-required megaplasmids pSymA (ranging from 93.3% to 100%; Fig. S4G, H and I), in particular in a specific region of 26 kb present in the genome of these three strains only (Fig. 3A). On the contrary, 60% of the detected orthologous gene hits in SM11 genomes were located on the chromid pSymB (Fig. S4J).

Based on COG annotations, the frequency of orthologous genes cluster hits associated with these k-mers was investigated (Table 2, Table S6). The distribution of candidate function of

gene hits within the most competitive strain genomes was not uniform. Enrichment for COG categories E (amino acid transport and metabolism), C (energy production and conversion), H (coenzyme transport and metabolism) and I (lipid transport and metabolism) was found in *S. meliloti* GR4, KH35c and KH46. Notably, the most represented orthologous genes groups are related to coenzyme F420 biosynthesis process (COG1060 and COG0391), transmembrane transport via ABC-type systems for branched-chain amino acids (COG0683, COG0410), and pyrimidine nucleotide biosynthetic process (COG0418) (Fig. 3B). Further, a putative caffeine dehydrogenase engaged in pathway of the caffeine transformation via C-8 oxidation (COG1529), and for the two subunits (PntA and PntB) of a presumptive proton-translocating NAD(P) transhydrogenase (COG3288, COG1282) liable for NADPH balancing mechanisms (Table 2) were found. Besides, other presumed functions common to the three strains (GR4, KH35c and KH46) were related to amino-acid degradation/catabolic process (COG1171, COG1788 and COG2057), carbohydrate metabolic process and oxidation-reduction process (COG2141), and transcriptional regulation by GntR-type regulator (COG2188) (Table 2).

The number of orthologous genes groups with functional annotation tagged by 51 best k-mers was lower in SM11. Except for orthologous group related to fatty acid metabolic process (COG1024), the candidate functions of gene hits identified in SM11 were exclusive. Among those, putative ABC-type ribose import-protein RbsA (COG1129), hypothetical autoinducer 2 import system permeases (COG1172), and a putative NAD-dependent glycerol dehydrogenase, belonging to large short-chain dehydrogenases/reductases family (SDR) (COG1028) were identified (Table 2).

Among the 103 CDSs tagged by 51 best k-mers, predicted protein-coding sequences (CDS) with no assigned function were also identified (Tables S5). A large part of these tagged CDS were identified in the SM11 genome and almost entirely located on the SM11 chromosome (Fig. S4A and F). In contrast, CDSs that were found in GR4, KH35c and KH46 genomes were located on homologs of the symbiosis-required megaplasmids pSymA (Fig. S4C, D and E). This last finding suggests the presence of a still remarkable amount of unknown genetic determinants potentially involved in the competition phenotype that require further investigation.

In the mapping procedure of k-mers related to the competition against BL225C, regulatory regions were also analyzed. We considered *bonafide* promoter sequences when hits mapped within 600 nucleotides upstream from CDS start, as previously reported (30). Ten of the 51 k-mers analyzed pinpointed 15 regulatory regions putatively associated with competition

phenotype in *S. meliloti* GR4, SM11, KH35c and KH46 (Table S7).

Seven regulatory region hits were associated with CDS with no assigned function (Table S7). These regulatory region hits were tracked exclusively in GR4 and SM11 genomes (Fig. S5A), and were mainly located on the chromosome of the two strains (Fig. S5C and D).

Eight regulatory region hits of putatively orthologous genes target were identified (Table S7, Fig. S5B). These eight regulatory region hits were entirely located on symbiosis-required megaplasmids pSymA in *S. meliloti* GR4, KH35c and KH46 (FigS5E, F and G) and on pSymB of *S. meliloti* SM11 (FigS5H).

In GR4, KH35c and KH46, the regulatory region hits are associated with genes encoding for proteins whose functions (COG1529 and COG2188) was previously observed (Fig. 3B and C), suggesting a role for genes in these COGs in determining competition phenotypes. In SM11, the regulatory regions hits are associated with a LysR-type orthologous gene (transcriptional regulation) and with a gene belonging to COG2308 possibly involved in the biosynthesis of small peptides.

DISCUSSION

Rhizobia-legume symbiosis is a paradigmatic example of bacteria-plant association, where bacteria behave as facultative symbionts, colonizing the plant tissue in nodules and establishing intracellular infection. The ability to colonize plant tissue is under selective pressure and rhizobial strains, which efficiently colonize host plants, have been shown to more effectively promote plant growth, giving rise to a partial “fitness alignment” between the host and the symbiont (33, 34). However, in nature, several strains compete for forming symbiotic associations with the host plant and often nodules are simultaneously colonized by different strains, which in turn may have different efficiency in promoting plant growth, some of them behaving as cheaters also (7, 8, 16, 35). Then, measuring the competitiveness for symbiosis and nodule colonization and predicting this phenotype from rhizobial genome sequencing is of paramount importance for understanding the evolution of rhizobia-plant symbiosis and developing effective inoculant strains for increasing agricultural yield of legume crops (8).

Here we addressed the possibility to develop a predictive model of strain competitiveness in the host plant–rhizobial symbiont systems *M. sativa*–*S. meliloti*. Direct measurements of competitiveness were obtained through the analysis of nodule occupancy. This experimental design pointed out a wide variety of strain response to the diverse competitive conditions,

identifying three different competition patterns and outlining a highly complex phenotype that strongly depends on the engaged competitor. Strains used in this work were originally isolated from different *Medicago* species, however they displayed good competition capabilities in nodulation of *M. sativa*, indicating that nodulation competitiveness is not strictly bound to the host genotype. Moreover, an abundant presence of mixed nodules was observed, confirming previous results (7, 11, 36). Mixed nodules were found in all the three sets of experiments suggesting that the possibility to co-infect nodules by different strains could be more widespread than expected. Strains M270 and T073, which were characterized by low nitrogen fixation rates, also showed low competition capabilities. Conversely, the differential response of strains with medium-high nitrogen-fixation efficiency advance the idea that a greater competitive ability is not correlated with high nitrogen fixation efficiency in *M. sativa–S. meliloti* interaction as previously suggested (37, 38). Concerning the competition against BL225C, except for GR4 and SM11, this assumption seems to be particularly true for highly efficient N-fixer strains (USDA1157, CCMM-B554, RU11/001, 2011, and AK58), which turned out to be medium-weak competitors. Co-persistence strategies in mixed infections based on the exploitation of benefits by less mutualistic strains were previously hypothesized (39). Indeed, for these strains, low competitive abilities coupled with a high percentage of mixed nodules hinted at the possibility that exploitation of more effective nitrogen fixer strains in mixed nodules by competitor may occur.

For the evaluation of the genetic determinants responsible of an increased competition phenotype, we used a method based on the use of k-mers, in GWAS, this kind of approaches are progressively increased in recent years (18, 40, 41), thanks to their flexibility in capturing different types of genetic variants and overcoming the alignment of sequence to a reference genome. PhenotypeSeeker is one of the up-to-date tools that utilize machine learning for predicting phenotypes from the sole genome sequences (32). This software combines the possibility to identify k-mers as signatures of competition phenotype with the training of a predictive linear regression model. Besides supplying statistical information that establishes the association degree between the presence of the k-mers in samples with their assessed competition phenotype, the evaluation of predicting power of built models can also be a hint of the degree of the association of top k-mers with phenotype. The results of the analyses with different combinations of training/test set ensure the robustness of the linear regression model. The influence exerted by different combinations of training/test set on the accuracy of the model, marked that competition phenotypes characterized by either positive or negative trends are less suitable for the regression analysis. In the case of Rm1021 dataset, the negative value

of averaged R^2 indicates that the mean of the phenotype values of training samples have more predictive power on the test set than the model itself. In contrast, the heterogeneous phenotypes ascertained in competition against BL225C might be well predicted by the trained model.

The evaluation of the predictive power of k-mers involved in model training is also a clue of the degree of the association of the latter to the phenotype competitiveness. Beyond the statistical analyses, the 51 k-mers related to the competition against BL225C that were taken into account are also highly predictive k-mers. Aside from having a significant impact on the outcomes of the machine-learning model, these k-mers, which featured a positive value of coefficients in the regression model, are positively correlated with the attested phenotype. Moreover, since they mapped in genomes of the four most competitive strains, these 51 k-mers can be considered as the most informative k-mers, allowing us to tag the genetic variants associated with remarkable competition capabilities.

The largest part of genes putatively associated with competitiveness was harboured by the megaplasmid pSymA (or pSymA homologs depending on the strain). Megaplasmid pSymA is the genomic element carrying all the genes necessary for symbiosis (e.g. *nod*, *fix* and *nif* genes) (42). Moreover, it was shown to harbour the largest part of genomic diversity in *S. meliloti* (43), then possibly largely contributing to the phenotypic diversity among strains and, on the basis of the obtained results, it could be linked to competition capabilities too.

Previous studies highlighted the importance of exopolysaccharide production, motility and signalling for symbiosis establishment and competition (25, 27). The genes (with functional annotation) found in this work, putatively associated with competitiveness, are mostly related to biosynthesis and transport functions. Many k-mers are related to COG1060 that, together with COG0391, is linked to the presence of the *fbi* operon in KH46, KH35c and GR4 strains only. The fbi operon is widely distributed in aerobic soil bacteria and it is responsible for the synthesis of the functional versatile redox factor F420 (44, 45). This cofactor is involved in the redox modification of many organic compounds facilitating low-potential two-electron redox reactions (44, 45). The presence of this cofactor is linked to several important processes such as persistence, antibiotic biosynthesis (tetracyclines, lincosamides and thiopeptides) and pro-drug activation (46); all functions that could play a central role in the competitiveness.

Another group of COGs highly represented in our dataset is related to ABC transporters. It is well known that *S. meliloti* genome encodes for a large number of ABC uptake and export systems (47, 48). This feature is probably linked to selective adaptation to oligotrophic soils and rhizospheric conditions (48, 49). According to our association analysis, a group of k-mers

tagged putative genes encoding for ATPase (COG0410) and permease (COG0683) subunits of branched-chain amino acids ABC transport complex Bra/Liv (50). In *S. meliloti*, a double mutant for the two main amino acid ABC transport complexes (aap bra) showed no reduction in N_2 fixation efficiency as well as no influence on the plant phenotype, suggesting the idea that in bacteroids a branched-chain amino acids auxotrophy, called "symbiotic auxotrophy", does not occur (51). However, an attenuated competitive phenotype was found in *S. meliloti* mutated in *livM* gene, which encodes for the permease subunit of Bra/Liv complex (52). It is then reasonable to suppose that this complex may provide a noteworthy benefit in the competition dynamics ensuring a higher supply of amino acids in free-living rhizospheric conditions and increasing strain competitiveness (52, 53). Other COGs related to ABC transporter are COG1129 and COG1172. Proteins grouping in COG1129 are ATPase component of an ABC-type ribose import system. In *Rhizobium leguminosarum*, a putative ribose ABC transporter (RbsA, RL2720) is induced by the presence of arabinogalactan, and it is specifically overexpressed in alfalfa rhizosphere (49). COG1172 contains ribose/xylose/arabinose/galactoside ABC-type transport system permease component remarking the importance of efficient carbon uptake in the rhizosphere to outcompete other bacteria. COG1172 is also related to the import of autoinducers signalling molecules in the quorum sensing process, whose connection with *S. meliloti* competitive behavior is well known (54).

Several COGs are related to the metabolism of different compounds: pyrimidine (COG0418; dihydroorotase), glutamate (COG1788 and COG2057; glutaconate CoA transferase), amino acids (COG1171; threonine dehydratase), fatty acids (COG1024; enoyl-CoA hydratase/carnitine racemase) and glycerol (COG1028; glycerol dehydrogenase), reinforcing the importance of metabolic versatility to increase adaptation to the rhizosphere. Indeed, in *R. leguminosarum* bv. *viciae* a deficiency for glycerol transport or utilization leads to reduced competitiveness for nodulation (55).

At last, a group of k-mers fell within transcriptional regulation genes (COG2188, COG1802 and COG0583) suggesting their possible involvement for a fine-tuning bacterial response to the presence of other competitors and/or for a quick response to variation of the external conditions.

Other COGs retrieved with our approach have less clear connections with competitiveness and will require further studies to infer their possible role in this process: COG1282/3288 (NAD(P) transhydrogenase) that may be involved in removing reactive oxygen species, COG0210 (helicase) and COG1529 (caffeine dehydrogenase).

A substantial part of k-mers maps on hypothetical genes with unknown functions, indeed many genes required for rhizobial adaptation to the rhizosphere are not characterized yet. Transcriptomic analysis of rhizobia isolated from the rhizosphere showed the expression of many hypothetical genes (49). This finding suggests that there is still to discover in the pangenome of *S. meliloti* a number of functions potentially important in the fitness associated with the symbiotic interaction and possibly in plant growth promotion.

Rhizobial competitiveness is a cornerstone for plant colonization; therefore, the selection of highly competitive/efficient rhizobia is fundamental for sustainable agricultural production. Here, we have reported on the feasibility and reliability of using a k-mer-based GWAS approach to detect genes associated with this complex quantitative phenotype in the plant symbiont *S. meliloti*. Several functions contribute to ameliorate competitiveness indicating that many different bricks, increasing rhizobial versatility, pave the way for success in competition. This approach may provide the basis for large-scale screening of putative competitiveness capabilities among pairs of strains on the basis of genome sequences. Interestingly, the evidence that most of the genes putatively associated with competition resides on the megaplasmid pSymA can offer the possibility to extend the creation of *ad hoc* hybrid strains by mobilizing the pSymA megaplasmid from different hosts (56) in the aim to develop novel ameliorated inoculants (8).

MATERIAL AND METHODS

Bacterial strains, plasmids and growth conditions. The strains and plasmids used in this work are listed in Table S7. Escherichia coli strains were grown in liquid or solid Luria Bertani (LB) medium (Sigma Aldrich) at 37°C (57), supplemented with tetracycline (10 µg/ml). *Sinorhizobium meliloti* strains were cultured in broth or agar tryptone yeast (TY) medium with 0.2 g/l CaCO3 at 30° C (58), supplemented with streptomycin (500 µg/ml in broth and agar media), rifampicin (50 µg/ml) and tetracycline (1 µg/ml in liquid broth medium, 2 µg/ml in agar medium), when necessary.

Competition assay. *Medicago sativa* (cv. Maravigliosa) plantlets were germinated and grown as described in nodulation assays section (Supplementary methods). *S. meliloti* strains were grown at 30° C to late exponential phase (OD_{600} = 0.6 - 0.8) in TY with opportune antibiotics; then each culture was washed 2 times in Nitrogen-free solution and diluted to a final concentration of $5x10^4$ CFU/ml. The inocula mixtures were prepared with equal volumes of

cellular suspensions of two different fluorescent-tagged strains. A total of 39 competition experiments were set up (13 GFP-tagged strains x 3 RFP-tagged strains). Six plants for each competition experiment were inoculated with 1 ml of inocula mixtures per seedling. After 4 weeks, nodule fluorescence was detected using a fluorescence stereomicroscope, Stereo Discovery V12 (Zeiss; Germany, Oberkochen), equipped with a CCD camera controlled by Axiovision software for image acquisition. All fresh nodules of each plant were individually exposed with filters for GFP (Zeiss Filter Set 38HE, excitation 470/40 emission 525/50) and DsRed (Zeiss Filter Set 43HE, excitation 550/25 and emission 605/70). The images obtained were processed with ImageJ software (59). The number of green nodules (occupied/produced by GFP-tagged strains), red nodules (occupied/produced by RFP-tagged strains) and mixed nodules (occupied/produced by both fluorescent-tagged strains) were counted for each plant of each competition experiment. The nodule occupancy was expressed as the ratio of the number of nodules (green, red or mixed) by the total number of nodules present on the roots of each plant. Statistical analyses of data were performed with nonparametric Kruskal-Wallis and Dunn test post-hoc by using FSA and rcompanion packages of Rstudio software (60).

PhenotypeSeeker analysis. The single nodule occupancy of the strains, assessed in the three competition experiments, were converted in a continuous matrix of equivalent values between 0 and 1 for each dataset. The FASTA genome sequences of 13 strains and obtained matrices were used as input to count all k-mers for each set of competition. The k-mer length was set to 13 nucleotides. In the first filtering step, the k-mers that were present in or missing from less than two samples ("—min 2—max 2"; default) were rejected. The clonal population structure correction was performed. The k-mers were tested for the analyses of association with the phenotype according to the weighted Welch two-sample t-test, and the k-mers with a p-value higher than 0.05 were automatically discarded. The linear regression models were achieved using only top lowest p-valued k-mers for all three datasets, with default parameters. For the first regression analyses, the strains were split into a training test (KH46, CCMM-B554, T073, Rm41, HM006, 2011, USDA1157, SM11 and RU11/001) and test set (AK58, KH35c, GR4, M270). In the second regression analyses, three different random combinations of training and test set were used, and the model evaluation metrics were averaged over 3-fold train/test splits. The 3-fold explicitly indicates that each strain was once included into test set and twice included into training set.

SUPPLEMENTAL MATERIAL

Supplementary methods

Fig. S1

Fig. S2

Fig. S3

Fig. S4

Fig. S5

Table S1

Table S2

Table S3

Table S4

Table S5

Table S6

Table S7

Table S8

ACKNOWLEDGEMENTS

Erki Aun and Maido Remm were funded by the institutional grant IUT34-11 from the Estonian Ministry of Education and Research and the EU ERDF grant No. 2014-2020.4.01.15-0012 (Estonian Center of Excellence in Genomics and Translational Medicine).

A.B., F.P., A.M. and C.V. conceived and planned the research; A.B. and F.P. performed competition assays and symbiotic assay; E.Az. and F.D. contributed to competition assays; E.Au. and M.R performed PhenotypeSeeker analysis and contributed to linear regression model interpretation; A.B. and G.B. performed the bioinformatics analyses; A.B., F.P., A.M, L.G., and C.V. interpreted data; A.B., F.P., A.M. prepared the manuscript; all authors participated in editing the manuscript.

We declare no competing interests.

REFERENCES

1. Masson-Boivin C, Sachs JL. 2018. Symbiotic nitrogen fixation by rhizobia — the roots of a success story. Curr Opin Plant Biol 44:7–15.

2. Masson-Boivin C, Giraud E, Perret X, Batut J. 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? Trends Microbiol 17:458–466.

3. Oldroyd GED, Murray JD, Poole PS, Downie JA. 2011. The rules of engagement in the legume-rhizobial symbiosis. Annu Rev Genet 45:119–144.

4. Haag AF, Arnold MFF, Myka KK, Kerscher B, Dall'Angelo S, Zanda M, Mergaert P, Ferguson GP. 2013. Molecular insights into bacteroid development during *Rhizobium*–legume symbiosis. FEMS Microbiol Rev 37:364–383.

5. Sprent JI. 2009. Legume nodulation: a global perspective. John Wiley & Sons.

6. Hirsch AM. 1992. Developmental biology of legume nodulation. New Phytol 122:211–237.

7. Checcucci A, Azzarello E, Bazzicalupo M, Galardini M, Lagomarsino A, Mancuso S, Marti L, Marzano MC, Mocali S, Squartini A, Zanardo M, Mengoni A. 2016. Mixed nodule infection in *Sinorhizobium meliloti–Medicago sativa* symbiosis suggest the presence of cheating behavior. Front Plant Sci 7:835.

8. Checcucci A, DiCenzo GC, Bazzicalupo M, Mengoni A. 2017. Trade, diplomacy, and warfare: The quest for elite rhizobia inoculant strains. Front Microbiol 8:2207.

9. Werner GDA, Strassmann JE, Ivens ABF, Engelmoer DJP, Verbruggen E, Queller DC, Noe R, Johnson NC, Hammerstein P, Kiers ET. 2014. Evolution of microbial markets. Proc Natl Acad Sci 111:1237–1244.

10. Kiers ET, Rousseau RA, West SA, Denison RF. 2003. Host sanctions and the legume – rhizobium mutualism. Nature 425:1095–1098.

11. Denison RF. 2000. Legume sanctions and the evolution of symbiotic cooperation by rhizobia. Am Nat 156:567–576.

12. Porter SS, Simms EL. 2014. Selection for cheating across disparate environments in the legume-rhizobium mutualism. Ecol Lett 17:1121–1129.

13. Kiers ET, Rousseau RA, Denison RF. 2006. Measured sanctions: Legume hosts detect quantitative variation in rhizobium cooperation and punish accordingly. Evol Ecol 1077–1086.

14. Gano-Cohen KA, Wendlandt CE, Stokes PJ, Blanton MA, Quides KW, Zomorrodian A, Adinata ES, Sachs JL. 2019. Interspecific conflict and the evolution of ineffective rhizobia. Ecol Lett 22:914–924.

15. Gano-Cohen KA, Wendlandt CE, Al Moussawi K, Stokes PJ, Quides KW, Weisberg AJ, Chang JH, Sachs JL. 2020. Recurrent mutualism breakdown events in a legume rhizobia metapopulation. Proc R Soc B Biol Sci 287:20192549.

16. Mendoza-Suárez MA, Geddes BA, Sánchez-Cañizares C, Ramírez-González RH, Kirchhelle C, Jorrin B, Poole PS. 2020. Optimizing *Rhizobium*-legume symbioses by simultaneous measurement of rhizobial competitiveness and N_2 fixation in nodules. Proc Natl Acad Sci 117:9822–9831.

17. Kirschner MW, Gerhart JC, Norton J. 2005. The Plausibility of Life. Yale University Press.

18. Lees JA, Vehkala M, Välimäki N, Harris SR, Chewapreecha C, Croucher NJ, Marttinen P, Davies MR, Steer AC, Tong SYC, Honkela A, Parkhill J, Bentley SD, Corander J. 2016. Sequence element enrichment analysis to determine the genetic basis of bacterial phenotypes. Nat Commun 7:12797.

19. Moradigaravand D, Palm M, Farewell A, Mustonen V, Warringer J, Parts L. 2018. Prediction of antibiotic resistance in *Escherichia coli* from large-scale pan-genome data. PLOS Comput Biol 14:e1006258.

20. Rahman A, Hallgrímsdóttir I, Eisen M, Pachter L. 2018. Association mapping from sequencing reads using k-mers. Elife 7:e32920.

21. Power RA, Parkhill J, de Oliveira T. 2017. Microbial genome-wide association studies: Lessons from human GWAS. Nat Rev Genet 18:41–50.

22. Kroll S, Agler MT, Kemen E. 2017. Genomic dissection of host–microbe and microbe–microbe interactions for advanced plant breeding. Curr Opin Plant Biol 36:71–78.

23. Beilsmith K, Thoen MPM, Brachi B, Gloss AD, Khan MH, Bergelson J. 2019. Genome-wide association studies on the phyllosphere microbiome: Embracing complexity in host–microbe interactions. Plant J 97:164–181.

24. Epstein B, Abou-Shanab RAI, Shamseldin A, Taylor MR, Guhlin J, Burghardt LT, Nelson M, Sadowsky MJ, Tiffin P. 2018. Genome-wide association analyses in the model *Rhizobium Ensifer meliloti*. MSphere 3:e00386-18.

25. Burghardt LT, Epstein B, Guhlin J, Nelson MS, Taylor MR, Young ND, Sadowsky MJ, Tiffin P. 2018. Select and resequence reveals relative fitness of bacteria in symbiotic and free-living environments. Proc Natl Acad Sci 115:201714246.

26. Burghardt LT. 2019. Evolving together, evolving apart: Measuring the fitness of rhizobial bacteria in and out of symbiosis with leguminous plants. New Phytol doi:10.1111/nph.16045.

27. Boivin S, Ait Lahmidi N, Sherlock D, Bonhomme M, Dijon D, Heulin-Gotty K, Le-Queré A, Pervent M, Tauzin M, Carlsson G, Jensen E, Journet EP, Lopez-Bellido R, Seidenglanz M, Marinkovic J, Colella S, Brunel B, Young P, Lepetit M. 2020. Host-specific competitiveness to form nodules in *Rhizobium leguminosarum* symbiovar *viciae*. New Phytol 226:555–568.

28. Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC. 2007. How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. Nat Rev Microbiol 5:619–633.

29. diCenzo GC, Zamani M, Checcucci A, Fondi M, Griffitts JS, Finan TM, Mengoni A. 2019. Multidisciplinary approaches for studying *Rhizobium* – legume symbioses. Can J Microbiol 33:1–33.

30. Galardini M, Mengoni A, Brilli M, Pini F, Fioravanti A, Lucas S, Lapidus A, Cheng J, Goodwin L, Pitluck S, Land M, Hauser L, Woike T, Mikhailova N, Ivanova N, Daligault H, Bruce D, Detter C, Tapia R, Han C, Teshima H, Mocali S, Bazzicalupo M, Biondi EG. 2011. Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. BMC Genomics 12:235.

31. Sugawara M, Epstein B, Badgley BD, Unno T, Xu L, Reese J, Gyaneshwar P, Denny R, Mudge J, Bharti AK, Farmer AD, May GD, Woodward JE, Médigue C, Vallenet D, Lajus A, Rouy Z, Martinez-Vaz B, Tiffin P, Young ND, Sadowsky MJ. 2013. Comparative genomics of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. Genome Biol 14:R17.

32. Aun E, Brauer A, Kisand V, Tenson T, Remm M. 2018. A k-mer-based method for the identification of phenotype-associated genomic biomarkers and predicting phenotypes of sequenced bacteria. PLoS Comput Biol 14:e1006434.

33. Friesen ML. 2012. Widespread fitness alignment in the legume-rhizobium symbiosis. New Phytol 194:1096–1111.

34. Westhoek A, Field E, Rehling F, Mulley G, Webb I, Poole PS, Turnbull LA. 2017. Policing the legume-*Rhizobium* symbiosis: A critical test of partner choice. Sci Rep 7:1419.

35. Kiers ET, Ratcliff WC, Denison RF. 2013. Single-strain inoculation may create spurious correlations between legume fitness and rhizobial fitness. New Phytol 198:4–6.

36. Gage DJ. 2002. Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*. J Bacteriol 184:7042–7046.

37. Amarger N. 1981. Competition for nodule formation between effective and ineffective strains of *Rhizobium meliloti*. Soil Biol Biochem 13:475–480.

38. Bourion V, Heulin-Gotty K, Aubert V, Tisseyre P, Chabert-Martinello M, Pervent M, Delaitre C, Vile D, Siol M, Duc G, Brunel B, Burstin J, Lepetit M. 2018. Co-inoculation of a pea core-collection with diverse rhizobial strains shows competitiveness for nodulation and efficiency of nitrogen fixation are distinct traits in the Interaction. Front Plant Sci 8:2249.

39. Friesen ML, Mathias A. 2010. Mixed infections may promote diversification of mutualistic symbionts: Why are there ineffective rhizobia? J Evol Biol 23:323–334.

40. Drouin A, Giguère S, Déraspe M, Marchand M, Tyers M, Loo VG, Bourgault AM, Laviolette F, Corbeil J. 2016. Predictive computational phenotyping and biomarker discovery using reference-free genome comparisons. BMC Genomics 17:754.

41. Davis JJ, Boisvert S, Brettin T, Kenyon RW, Mao C, Olson R, Overbeek R, Santerre J, Shukla M, Wattam AR, Will R, Xia F, Stevens R. 2016. Antimicrobial resistance prediction in PATRIC and RAST. Sci Rep 6:27930.

42. Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, Bothe G, Boutry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Dreano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernandez-Lucas I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetelle D, Purnelle B, Ramsperger U, Surzycki R, Thebault P, Vandenbol M, Vorholter FJ, Weidner S, Wells DH, Wong K, Yeh KC, Batut J. 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science (80) 293:668–672.

43. Galardini M, Pini F, Bazzicalupo M, Biondi EG, Mengoni A. 2013. Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*. Genome Biol Evol 5:542–558.

44. Ney B, Ahmed FH, Carere CR, Biswas A, Warden AC, Morales SE, Pandey G, Watt SJ, Oakeshott JG, Taylor MC, Stott MB, Jackson CJ, Greening C. 2017. The methanogenic redox cofactor F420 is widely synthesized by aerobic soil bacteria. ISME J 11:125–137.

45. Greening C, Ahmed FH, Mohamed AE, Lee BM, Pandey G, Warden AC, Scott C, Oakeshott JG, Taylor MC, Jackson CJ. 2016. Physiology, biochemistry, and applications of F420 - and Fo -dependent redox reactions. Microbiol Mol Biol Rev 80:451–493.

46. Bashiri G, Antoney J, Jirgis ENM, Shah M V., Ney B, Copp J, Stuteley SM, Sreebhavan S, Palmer B, Middleditch M, Tokuriki N, Greening C, Scott C, Baker EN, Jackson CJ. 2019. A revised biosynthetic pathway for the cofactor F 420 in prokaryotes. Nat Commun 10:1558.

47. Torres-Quesada O, Millán V, Nisa-Martínez R, Bardou F, Crespi M, Toro N, Jiménez-Zurdo JI. 2013. Independent activity of the homologous Small regulatory RNAs AbcR1 and AbcR2 in the legume symbiont *Sinorhizobium meliloti*. PLoS One 8:e68147.

48. Mauchline TH, Fowler JE, East AK, Sartor AL, Zaheer R, Hosie AH, Poole PS, Finan TM. 2006. Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. Proc Natl Acad Sci U S A 103:17933–17938.

49. Ramachandran VK, East AK, Karunakaran R, Downie JA, Poole PS. 2011. Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biol 12:R106.

50. Hosie AH, Allaway D, Galloway CS, Dunsby HA, Poole PS. 2002. *Rhizobium leguminosarum* has a second general amino acid permease with unusually broad substrate specificity and high similarity to branched-chain amino acid transporters (Bra/LIV) of the ABC family. J Bacteriol 184:4071–4080.

51. Prell J, Bourdes A, Kumar S, Lodwig E, Hosie A, Kinghorn S, White J, Poole P. 2010. Role of symbiotic auxotrophy in the *Rhizobium*-legume symbioses. PLoS One 5:e13933.

52. Pobigaylo N, Szymczak S, Nattkemper TW, Becker A. 2008. Identification of genes relevant to symbiosis and competitiveness in *Sinorhizobium meliloti* using signature-tagged mutants. Mol Plant Microbe Interact 21:219–231.

53. van Dillewijn P, Soto MJ, Villadas PJ, Toro N. 2001. Construction and environmental release of a *Sinorhizobium meliloti* strain genetically modified to be more competitive for alfalfa nodulation. Appl Environ Microbiol 67:3860–3865.

54. McIntosh M, Krol E, Becker A. 2008. Competitive and Cooperative Effects in Quorum-Sensing-Regulated Galactoglucan Biosynthesis in *Sinorhizobium meliloti*. J Bacteriol 190:5308–5317.

55. Ding H, Yip CB, Geddes BA, Oresnik IJ, Hynes MF. 2012. Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. Microbiology 158:1369–1378.

56. Alice C, George D, Veronica G, Marco B, Anke B, Francesca D, Johannes D, Camilla F, M FT, Marco F, Claudio L, Paola T, Tiziano V, Carlo V, Alessio M. 2018. Creation and characterization of a genomically hybrid strain in the nitrogen-fixing symbiotic bacterium *Sinorhizobium meliloti*. ACS Synth Biol 7:2365–2378.

57. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

58. Beringer JE. 1974. R factor transfer in *Rhizobium leguminosarum*. J Gen Microbiol 84:188–198.

59. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675.

60. R Development Core Team. 2011. R: A language and environment for statistical computing. R Found Stat Comput.

Figures

Figure 1. Competition performances and epifluorescence stereomicroscope images. Bar plots showing the percentage of nodule occupancy of 13 *S. meliloti* strains in three sets of competition experiments: A) competition against *S. meliloti* Rm1021, B) competition against *S. meliloti* AK83 and C) competition against *S. meliloti* BL225C. Green bars represent single nodule occupancy of the strains tested whose ID is reported on the x-axis, in yellow is showed the percentage of mixed nodules (nodules occupied by both strains) and in red the single nodule occupancy of the competitor used in each set of experiments. Pictures show nodules of *M. sativa* inoculated with a mix of *S. meliloti* 1021 RFP-tagged and D) KH46 GFP-tagged or E) GR4 GFP-tagged; F) and H) *S. meliloti* AK83 RFP-tagged and HM006 GFP-tagged; and *S. meliloti* BL225C RFP-tagged and I) CCMM B554 GFP-tagged or J) RU11-001 GFP-tagged.

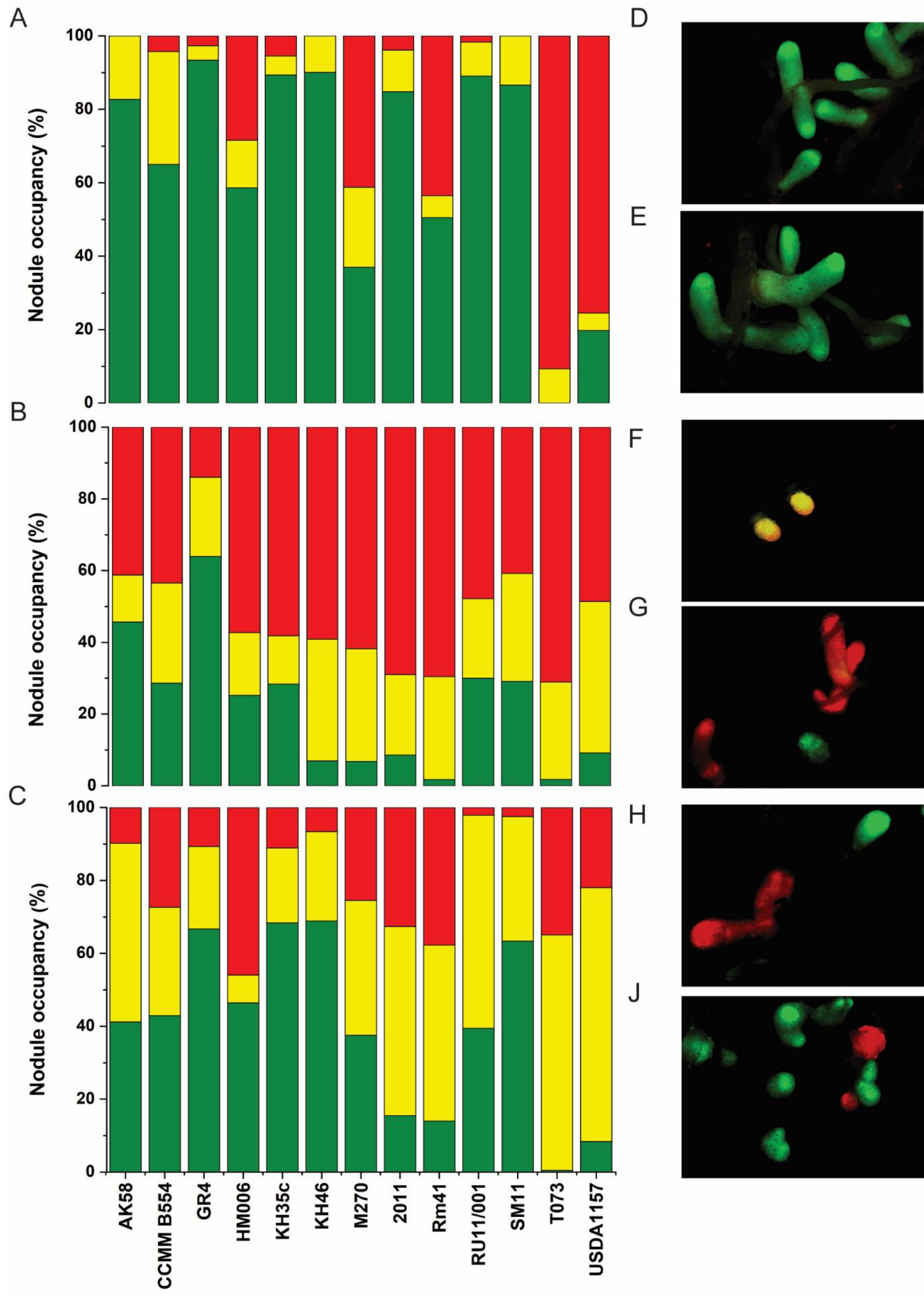
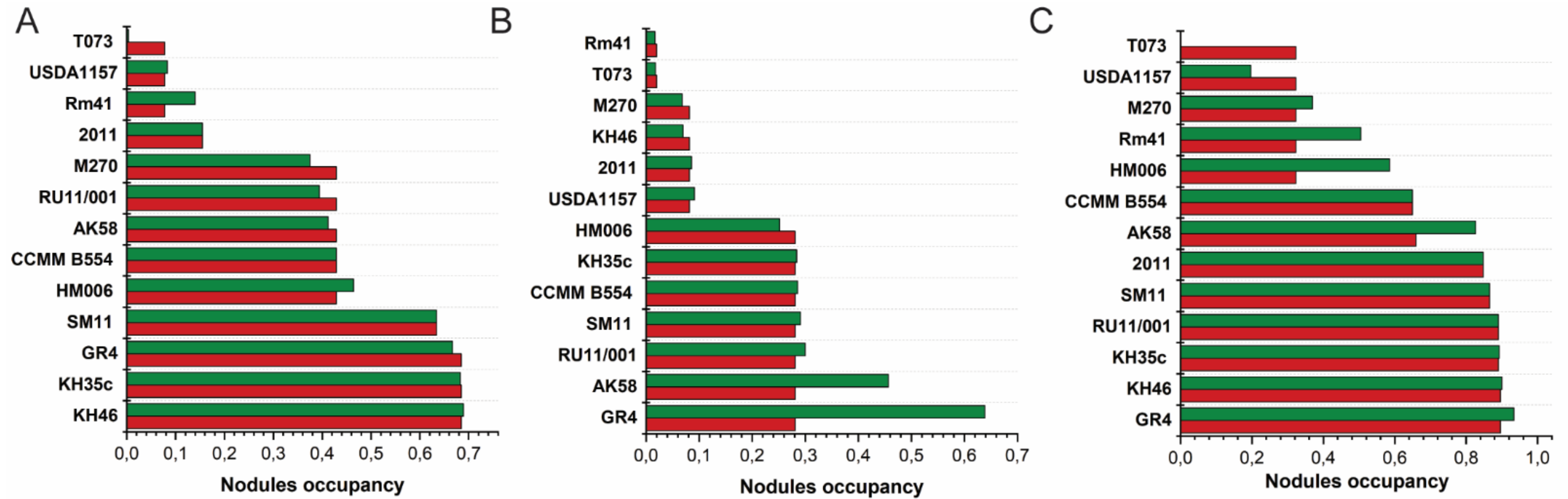
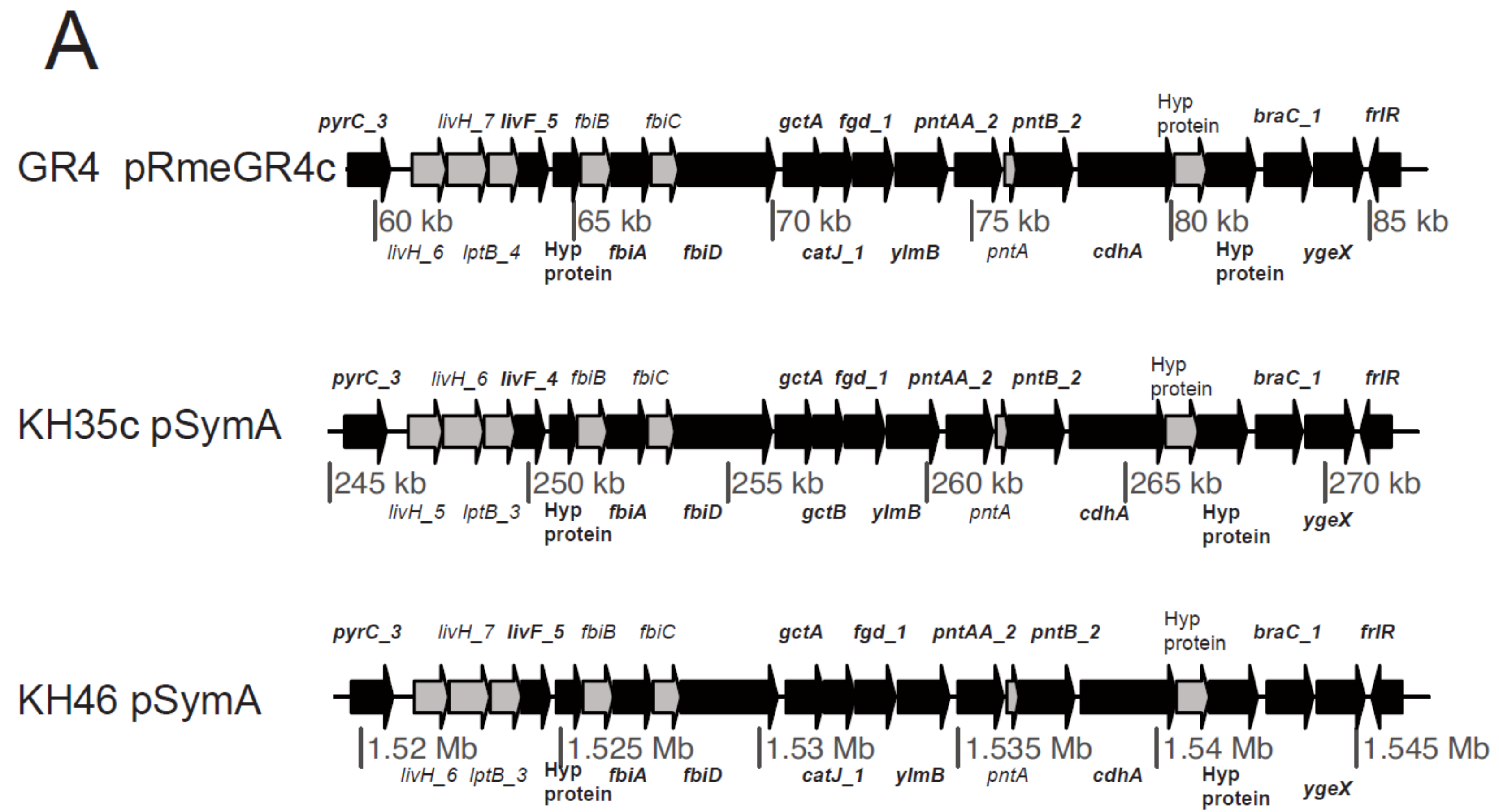

A
B
C
D
E
F
G
H
J
Nodule occupancy (%)
0
20
40
60
80
100
AK58
CCMM B554
GR4
HM006
KH35c
KH46
M270
2011
Rm41
RU11/001
SM11
T073
USDA1157

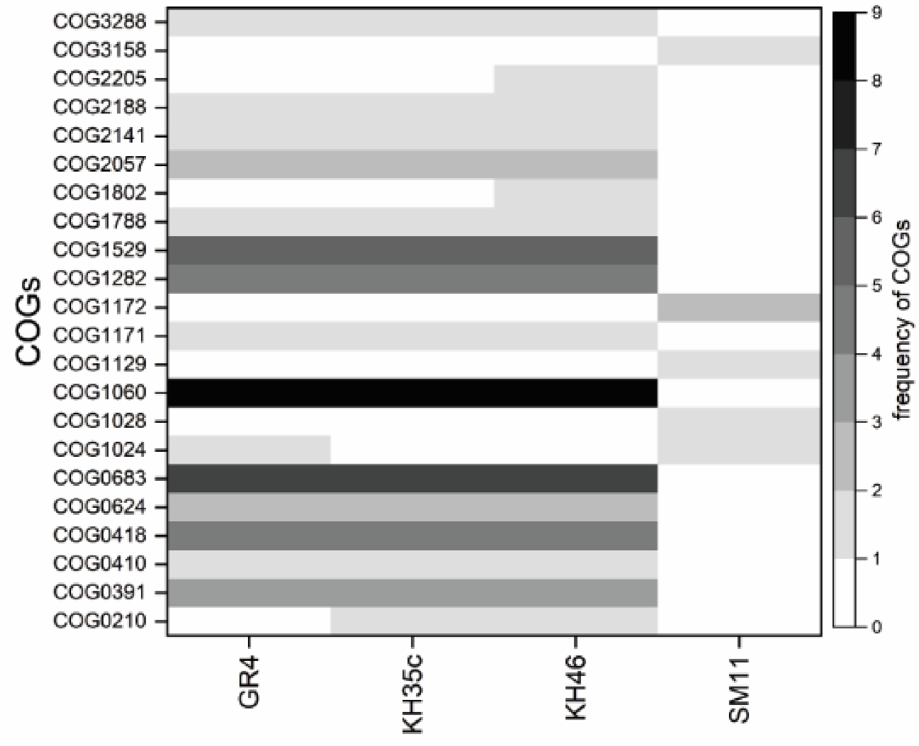
Figure 2. Predicted and actual competition patterns. Single nodule occupancy of *S. meliloti* strains in competition experiments versus *S. meliloti* strains A) Rm1021, B), AK83 and C) BL225C. Green bars indicate the actual phenotypes, red bars the predicted phenotypes

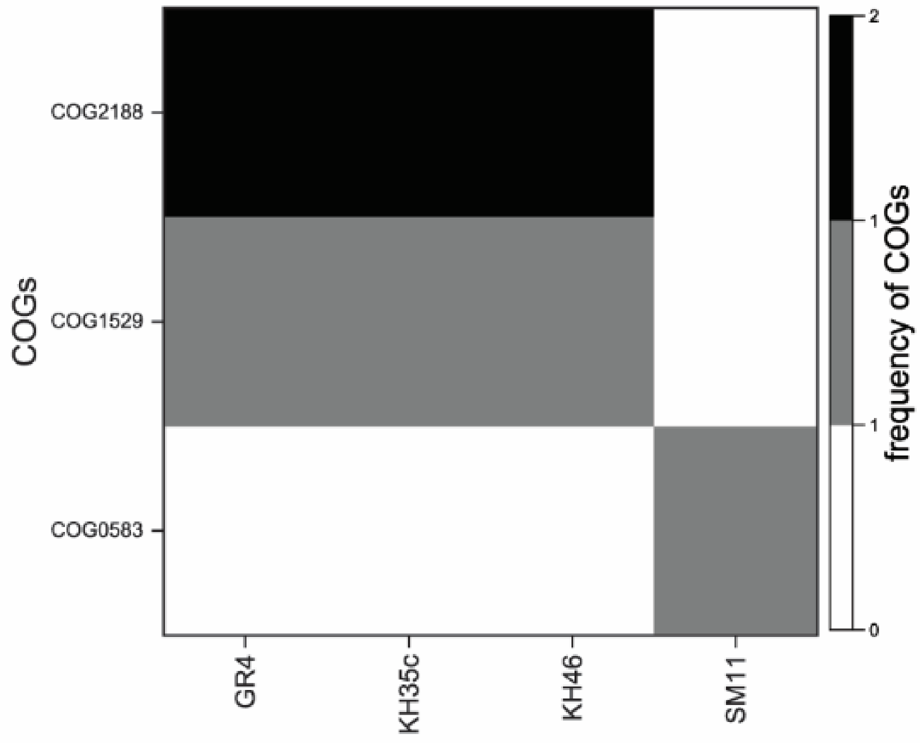


.

Figure 3. Genetic determinants associated with an increase of competition capabilities A) k-mers mapping in a region of the symbiotic megaplasmid (pSymA or homologue plasmids) present exclusively in the genome of *S. meliloti* GR4, KH35c and KH46; Genes containing one or more k-mers are indicated with a black arrow, gene annotation is referred to Prokka output. Frequency of candidate functions of B) gene hits and C) regulatory regions identified by 51 best k-mers in the highest competitive strains. The frequency of candidate functions reported as COG annotations (rows) in each strain (columns) is represented by grey-scale shades.




B
COGs
COG3288
COG3158
COG2205
COG2188
COG2141
COG2057
COG1802
COG1788
COG1529
COG1282
COG1172
COG1171
COG1129
COG1060
COG1028
COG1024
COG0683
COG0624
COG0418
COG0410
COG0391
COG0210
GR4
KH35c
KH46
SM11
frequency of COGs
9
8
7
6
5
4
3
2
1
0


C
COGs
COG2188
COG1529
COG0583
GR4
KH35c
KH46
SM11
frequency of COGs
2
1
1
0

Tables

Table 1. Performances of linear regression models for the three competition experiments. The results of both training and test set by PhenotypeSeeker are reported.

Dataset		The mean squared error	The coefficient of determination (R^2)	The Pearson correlation and p-value	The Spearman correlation coefficient and p-value	Total K-mers (p-value < 0.05)	Range p-value
Vs Rm1021	Training set	0.02	0.74	0.86, 0.0	0.98, 0.0	439886	4.99E-0^2 – 4.31E-0^3
	Test set	0.01	0.84	0.96, 0.04	1.0, 0.0		
Vs BL225C	Training set	0.0	0.98	0.99, 0.0	0.97, 0.0	182804	4.98E-0^2 – 1.35E-0^5
	Test set	0.0	0.96	0.99, 0.01	0.89, 0.11		
Vs AK83	Training set	0.0	0.99	0.99, 0.0	0.92, 0.0	292884	4.97E-0^2 – 1.05E-0^3
	Test set	0.04	0.11	0.8, 0.2	0.77, 0.23		

Table 2. List of functions putatively involved in promoting competing abilities. COG description of gene hits identified by 51 k-mers (p-value 1.13E-0[4]) in the most competitive strains (GR4, KH35c, KH46, SM11). Function/annotation are reported according to the annotation performed with Prokka in this work and/or using original annotation.

COG ID	COG classes	COG functional category	Prokka annotation / Other annotation	Biological process
COG0210	L	Superfamily I DNA or RNA helicase	ATP-dependent DNA helicase PcrA	Mismatch repair, Nucleotide excision repair
COG0391	GH	Archaeal 2-phospho-L-lactate transferase/Bacterial gluconeogenesis factor, CofD/UPF0052 family	Putative phosphoenolpyruvate transferase FbiA	Coenzyme F420 biosynthesis
COG0410	E	ABC-type branched-chain amino acid transport system, ATPase component	High-affinity branched-chain amino acid transport ATP-binding protein LivF	High-affinity branched-chain amino acid transport
COG0418	F	Dihydroorotase	Dihydroorotase PyrC	Pyrimidine nucleotide biosynthetis
COG0624	E	Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase or related deacylase	Probable N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine deformylase	Unknown function
COG0683	E	ABC-type branched-chain amino acid transport system, periplasmic component	Leucine-, isoleucine-, valine-, threonine- and alanine-binding protein Brac	Branched-chain amino acid transport
COG1024	I	Enoyl-CoA hydratase/carnithine racemase	Putaive fatty acid oxidation complex subunit alpha - enoyl-CoA hydratase FadJ	Fatty acid metabolism
COG1028	IQR	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family	Putative NAD-dependent glycerol dehydrogenase	Glycerol metabolic process
COG1060	H	2-iminoacetate synthase ThiH (tyrosine cleavage enzyme, thamine biosynthesis)	FbiC FO synthase	Coenzyme F0 biosynthesis

COG1129	G	ABC-type sugar transport system, ATPase component	Ribose import ATP-binding protein RbsA, CUT2 family	Ribose transmembrane transport
COG1171	E	Threonine dehydratase	Diaminopropionate ammonia-lyase	Cellular amino acid catabolic process
COG1172	G	Ribose/xylose/arabinose/galactoside ABC-type transport system, permease component	Autoinducer 2 import system permease protein LsrD	AI-2 transport system
COG1282	C	NAD/NADP transhydrogenase beta subunit	NAD(P) transhydrogenase subunit beta PntB	Oxidation-reduction process Nicotinate and nicotinamide metabolism
COG1529	C	CO or xanthine dehydrogenase, Mo-binding subunit	Putative caffeine dehydrogenase subunit alpha	Oxidation-reduction process
COG1788	I	Acyl CoA:acetate/3-ketoacid CoA transferase, alpha subunit	Glutaconate CoA-transferase GctA subunit A	Glutamate catabolic process (via hydroxyglutarate)
COG1802	K	DNA-binding transcriptional regulator, GntR family	HTH-type transcriptional repressor RspR	DNA-binding transcriptional regulation
COG2057	I	Acyl CoA:acetate/3-ketoacid CoA transferase, beta subunit	Glutaconate CoA-transferase GctB subunit B	Glutamate catabolic process (via hydroxyglutarate)
COG2141	HR	Flavin-dependent oxidoreductase, luciferase family (includes alkanesulfonate monooxygenase SsuD and methylene tetrahydromethanopterin reductase)	F420-dependent glucose-6-phosphate dehydrogenase	Carbohydrate metabolic process, oxidation-reduction process
COG2188	K	DNA-binding transcriptional regulator, GntR family	Putative transcriptional repressor	DNA-binding transcriptional regulation
COG2205	T	K+-sensing histidine kinase KdpD	two-component system sensor histidine kinase KdpD	Two-component regulatory system K+ sensing that regulate kdpABCF operon for Potassium transport
COG3158	P	K+ transporter	Low affinity potassium transport system protein Kup	Potassium ion transport
COG3288	C	NAD/NADP transhydrogenase alpha subunit	NAD(P) transhydrogenase subunit alpha PntA	Oxidation-reduction process Nicotinate and nicotinamide metabolism

Table 3. List of regulatory regions putatively involved in promoting competing capabilities. COG description of genes linked to regulatory region hits identified by 10 of the 51 k-mers associated to the most competitive strains (*S. meliloti* GR4, KH35c, KH46, SM11).

COG ID	COG classes	COG functional category	Prokka annotation / Product	Biological process
COG2308	S	Uncharacterized conserved protein, circularly permuted ATPgrasp superfamily	Uncharacterized putative protein	Function unknown
COG2188	K	DNA-binding transcriptional regulator, GntR family	Putative transcriptional repressor	DNA-binding transcriptional regulation
COG0583	K	DNA-binding transcriptional regulator, LysR family	HTH-type transcriptional regulator DmlR	Transcriptional regulator of dmlA (aerobic growth on D-malate as the sole carbon source)
COG1529	C	CO or xanthine dehydrogenase, Mo-binding subunit	Putative caffeine dehydrogenase subunit alpha	Oxidation-reduction process

SUPPLEMENTARY MATERIAL

Supplementary methods

Construction of *S. meliloti* fluorescently tagged strains. *S. meliloti* strains were tagged with green fluorescent protein (GFP) or red fluorescent protein (RFP). Donor *E. coli* S17-1 strains containing plasmids pHC60 (harboring a constitutively expressed GFP; (1) or pBHR mRFP (harboring a constitutively expressed RFP; (2) were used for biparental conjugations with rifampicin-resistant derivatives *S. meliloti* strains. Spontaneous rifampicin derivative *S. meliloti* strains were isolated by plating aliquots of 100µl of cell suspension of 109 cells on agar TY medium with rifampicin (50 µg/ml). Conjugal transfer was performed as previously described (3)

Nodulation and acetylene reduction assays. *Medicago sativa* (cv. Maravigliosa) seedlings were surface sterilized with 70% ethanol for 1 min, rinsed with sterile ddH_2O, treated with 2.5% sodium hypochlorite for 5 min and washed 20 times with sterile ddH_2O. Sterilized seeds were then let germinate on the cover of sterile plastic Petri dishes upside down for 4 days in the dark at room temperature. Seedlings were transferred in plastic pots containing a sterilized mixture of sand and vermiculite (ratio 2:3) and supplied with 120 ml of sterilized Nitrogen-free solution (1mM $CaCl_2$ $2H_20$, 0.1 mM KCl, 0.8 mM $MgSO_4$ $7H_2O$, 10 µM Fe EDTA, 35 µM H_3BO_3, 9 µM $MnCl_2$ $4H_2O$, 0.8 µM $ZnCl_2$, 0.5 µM Na_2MoO_4 $2H_2O$, 0.3 µM $CuSO_4$ $5H_2O$, 3.68 mM KH_2PO_4, 4 mM Na_2HPO_4 pH=6.5) (4). Seedlings were grown for 3 additional days before inoculation with *S. meliloti* strains. The strains were grown at 30°C to late exponential phase (OD_{600} < 0.6 - 0.8), washed 2 times in Nitrogen-free solution and then adjusted to an OD_{600} = 0.05 in Nitrogen-free solution. Nine plants for strains were inoculated with aliquots of 500µl of cell suspension of 5×10^7 CFU/ml, and grown in a growth chamber maintained at 23°C with a 16-h photoperiod. The same amount of Nitrogen-free solution was added to negative control plants (C^-). After 28 days, the epicotile length, number of nodules and dry weight were measured. For the acetylene-reduction assay, *M. sativa* plants were grown as described above. After 28 days, plants were collected in 100ml glass flasks (3 plants/flask) and sealed with gas-thigh silicone caps. Aliquots of 10ml of acetylene were injected into the flasks and, after 40 min, the ethylene concentration was measured by using a 7890B gas chromatograph system (Agilent technologies; California, USA), equipped with a 5975 Mass selective detector. Chromatographic analyses were performed in the following conditions: initial temperature, 40°C (isocratic for 10 min), gas flow (helium) 4 ml/min, injection 500µl (gas syringe) at a split ratio of 5:1. Nitrogen fixation rates were expressed in nanomoles of produced ethylene per hour,

per plant.

Statistical analysis of data was performed with Rstudio software (5). Shapiro test was performed to evaluate data distribution; ANOVA and Tukey post-hoc test or nonparametric Kruskal-Wallis and Dunn test post-hoc were performed using FSA and rcompanion packages.

Annotation and phylogenetic analyses. *S. meliloti* genomes were retrieved from the NCBI Genome Database (GenBank codes are reported in Table S1). Genome annotation of 13 *S. meliloti* strains was completed using Prokka (version 1.13) bacterial genome annotation tool (6). The pangenome of the 13 *S. meliloti* strains was constructed with Roary 3.11.3 (7) using default settings to construct a whole-genome phylogeny. Core genes alignment, obtained with Roary, was used to infer the evolutionary relationship of the strains tested as competitors. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site (8). The evolutionary history was reconstructed using the UPGMA method (bootstrap test of 1000 replicates). All ambiguous positions were removed for each sequence pair (pairwise deletion option). All evolutionary analyses were conducted using MEGA X software (9).

Mapping procedure. For each competing strain tested, the genome position of k-mers associated with the phenotype (competition against BL225C strain) was detected using the R package Biostrings (version 2.54) (10). Only k-mers aligning without mismatches or gaps on the positive or negative strand of the reference were taken into account to reflect the pipeline used by PhenotypeSeeker, which does not allow for mismatches. Absolute positions of k-mers were then transformed into relative ones based on genomic annotations following four rules:

1. If a k-mer was mapped inside a gene, its position was set to 0 independently from the strand

2. If a k-mer was mapped outside a gene on the positive strand, its position was adjusted by subtracting the starting position of the nearest gene. Since the starting position of the nearest gene on the plus strand is always greater than the starting position of the k-mer, the relative position will always be a negative value representing the number of bases ahead of the sequence of the gene on the reference genome.

3. If a k-mer was mapped outside a gene on the negative strand, its position was calculated by subtracting the starting position of the k-mer to the ending position of the gene. Analogously to the previous calculation, the starting position of the k-mer will always be greater than the ending position of the gene on the minus strand, thus the relative position of the k-mer will always be a negative value representing the number of bases behind the sequence of the gene on the

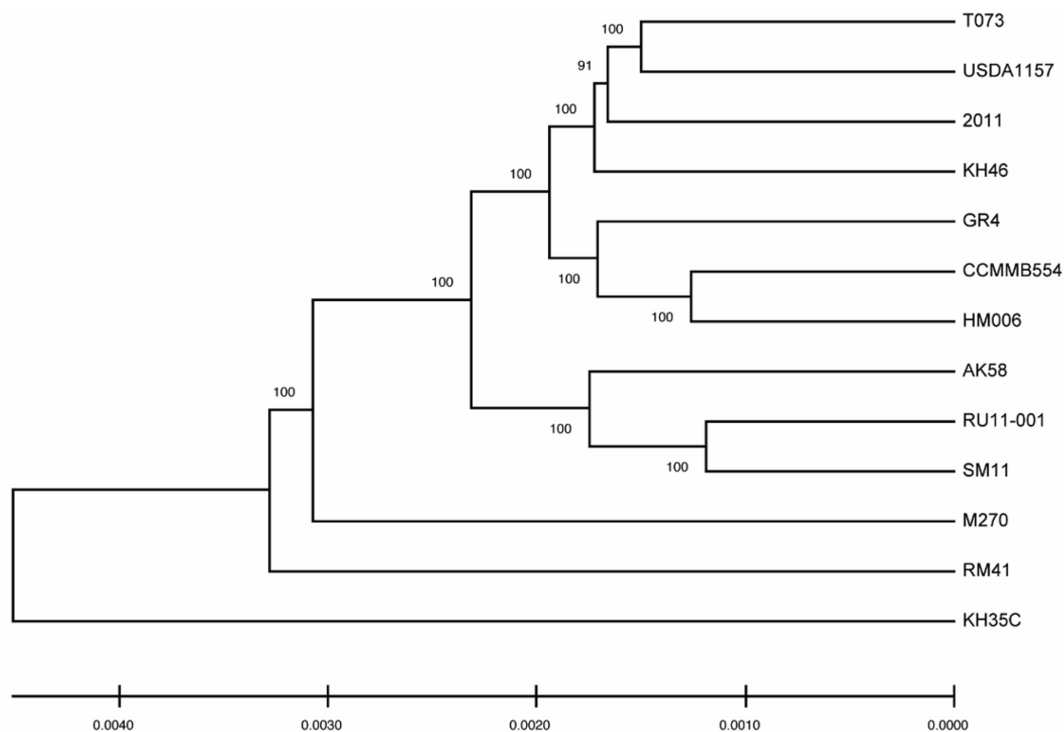
reference genome.

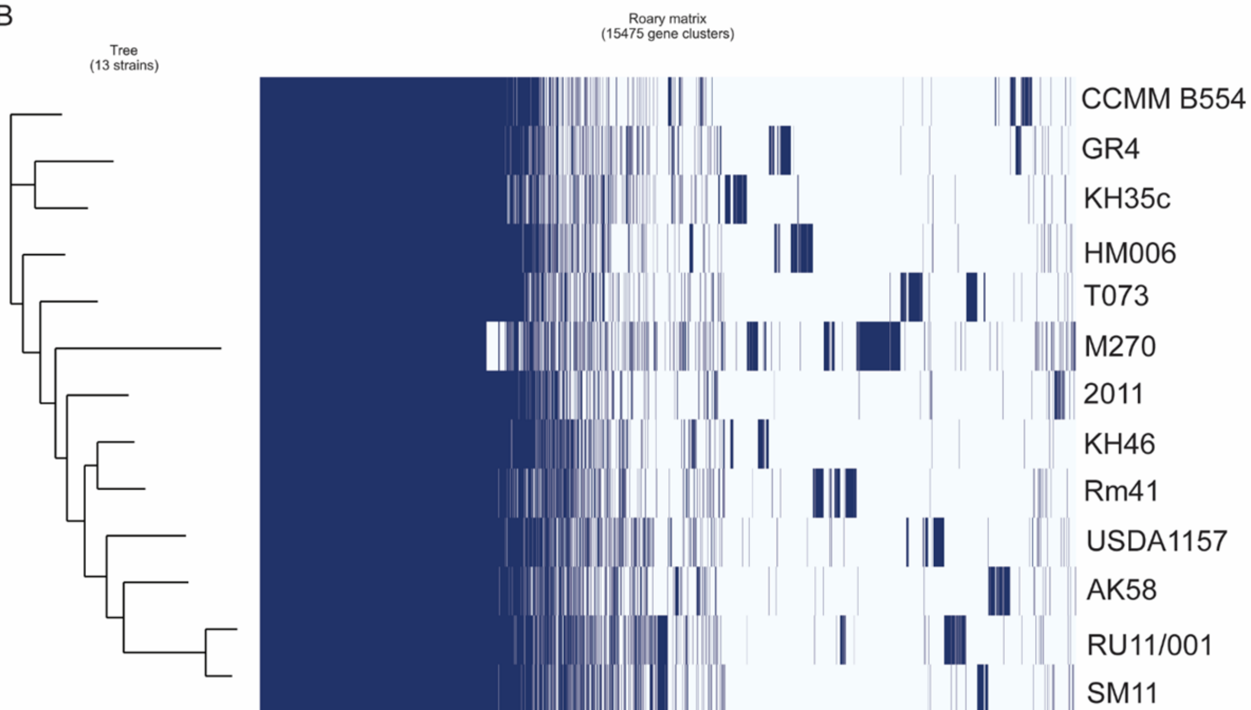
4. If a k-mer was mapped ahead of a gene on the positive strand and behind a gene on the negative strand (namely “between” two genes oriented in different directions), its position was calculated as reported in 1 and 2. Since both relative positions may be valid, they were both reported and considered in downstream analyses.

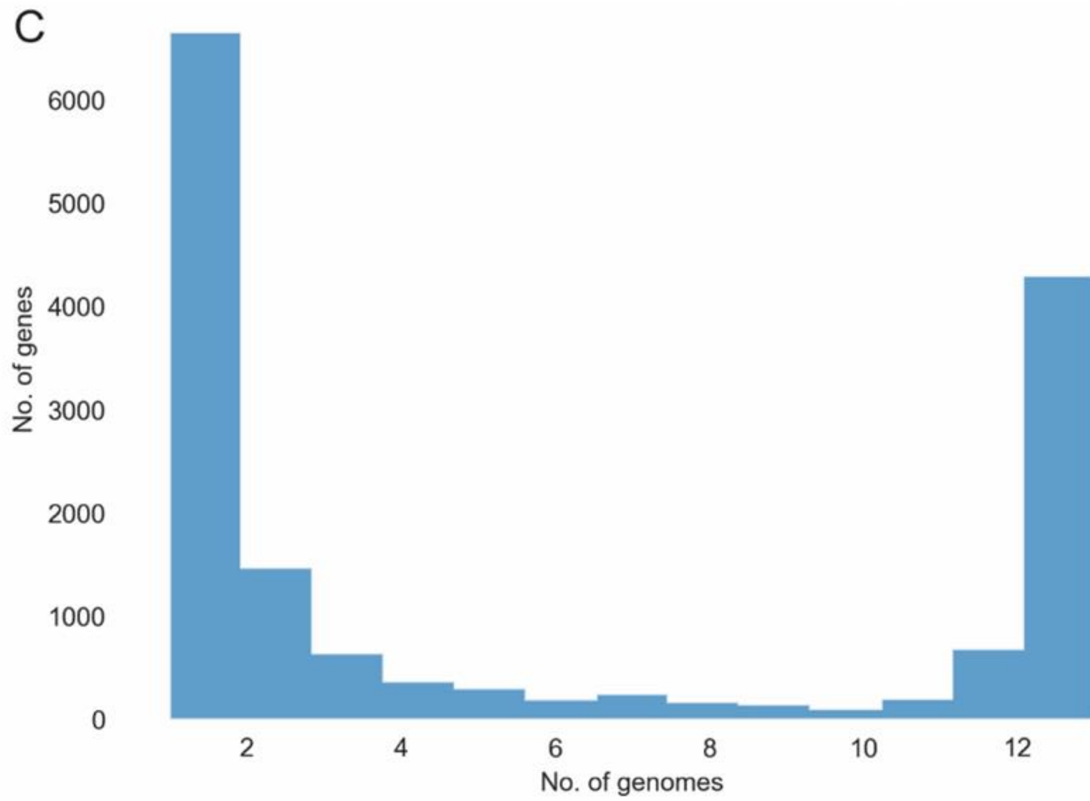
Relative position obtained were then used to extract the predicted protein-coding sequences (CDS) and regulatory regions mapped by 51 k-mers (with a p-value =1.31 e-0^4) by selecting those with a relative position equal to 0 and higher than -600 respectively.

Supplementary figures

Figure S1. Evolutionary relationships and pangenome of strains used as competitors. A) The evolutionary history was inferred using the UPGMA method on core genes concatenamer alignment. The optimal tree with the sum of branch length = 0.02977589 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10000 replicates) is shown next to the branches. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. There was a total of 3998704 positions in the final dataset. B) Heatmap showing gene presence (dark blue) or absence light blue in each strain, the tree on the left was build on the basis of presence/absence of genes. C) Histogram showing the frequency of genes depending on the number of genomes. D) Pie chart displaying all the genes present in the pangenome and their breakdown in the different genomes


A
T073
USDA1157
2011
KH46
GR4
CCMMB554
HM006
AK58
RU11-001
SM11
M270
RM41
KH35C
0.0040
0.0030
0.0020
0.0010
0.0000


B
Tree
(13 strains)
Roary matrix
(15475 gene clusters)
CCMM B554
GR4
KH35c
HM006
T073
M270
2011
KH46
Rm41
USDA1157
AK58
RU11/001
SM11


C
6000
5000
4000
3000
2000
1000
0
No. of genes
2
4
6
8
10
12
No. of genomes

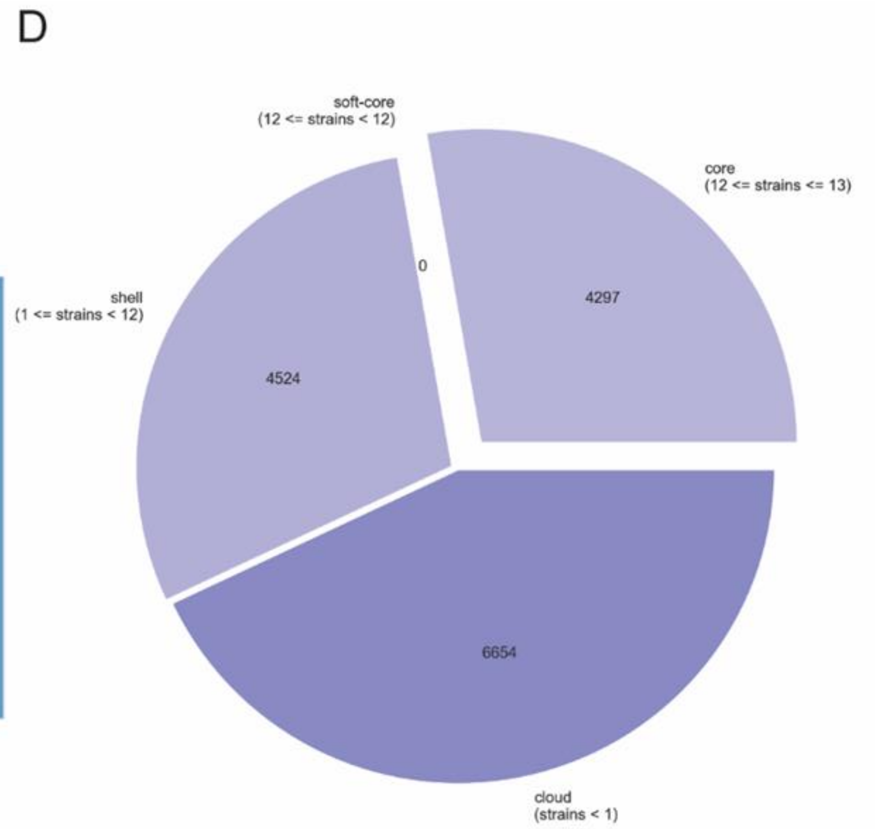
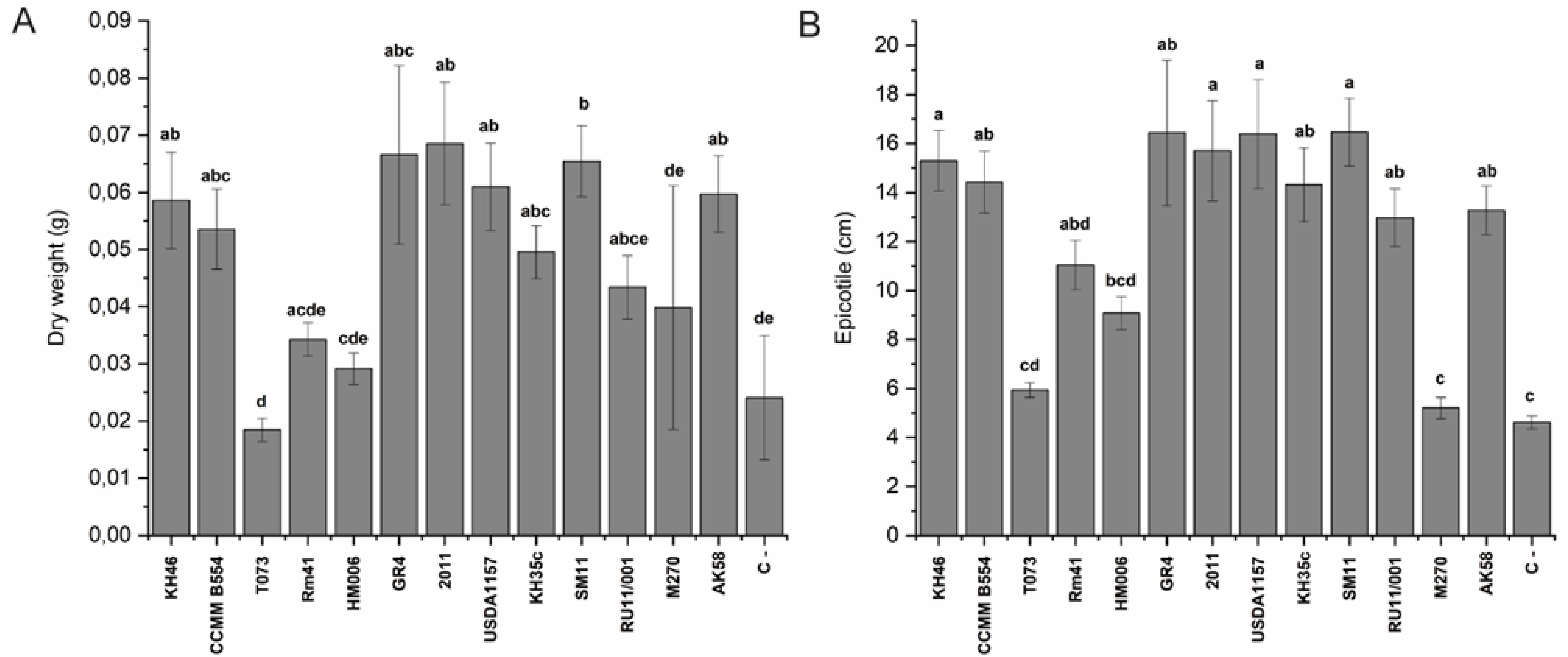
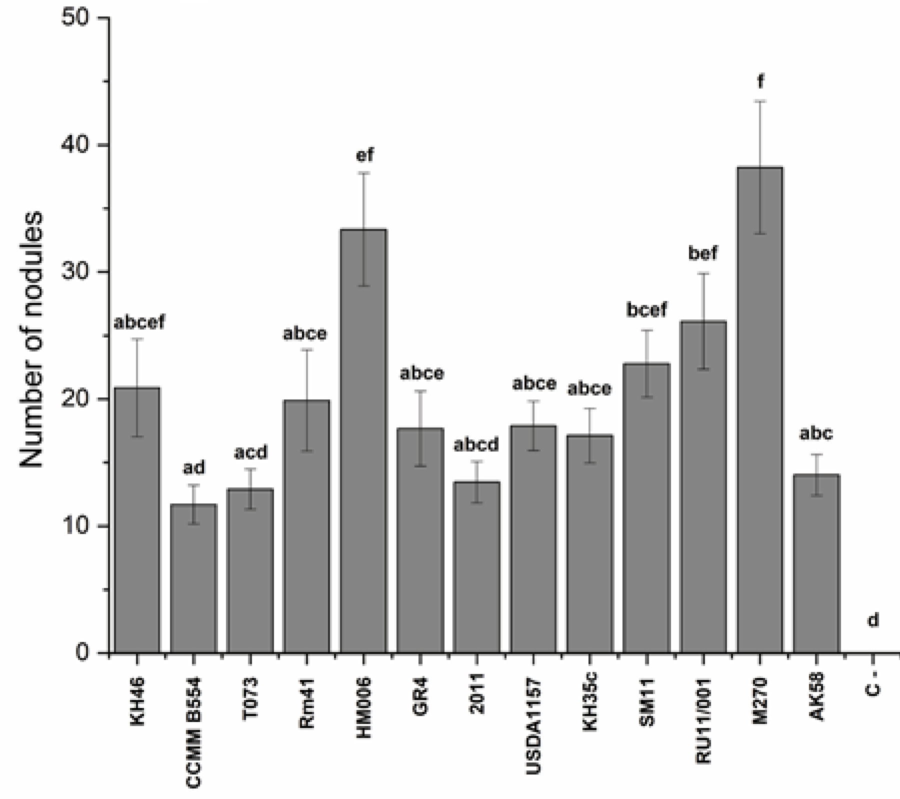

D
soft-core
(12 <= strains < 12)
core
(12 <= strains <= 13)
0
4297
shell
(1 <= strains < 12)
4524
6654
cloud
(strains < 1)

Figure S2. Nodulation assay and nitrogen fixation efficiency with single strains. a) Number of nodules per plant; b) Epicotyl length; c) Plant dry weight; d) Acetylene reduction assay (ARA).




C
Number of nodules
50
40
30
20
10
0
abcef
ad
acd
abce
ef
abce
abcd
abce
abce
bcef
bef
f
abc
d
KH46
CCMM B554
T073
Rm41
HM006
GR4
2011
USDA1157
KH35c
SM11
RU11/001
M270
AK58
C -

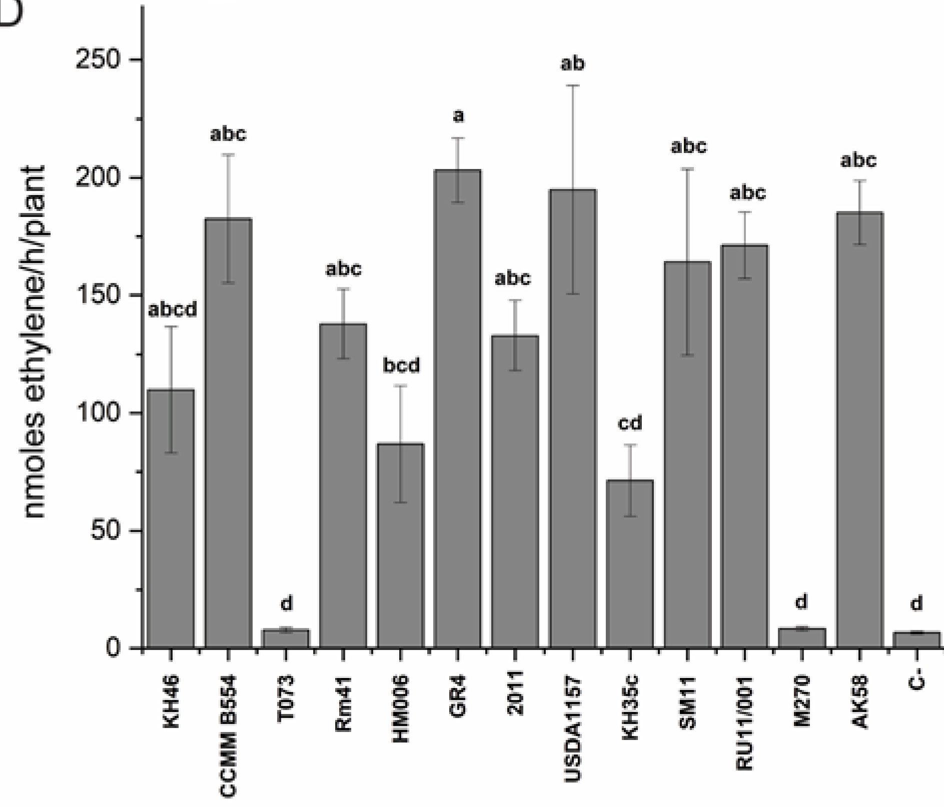
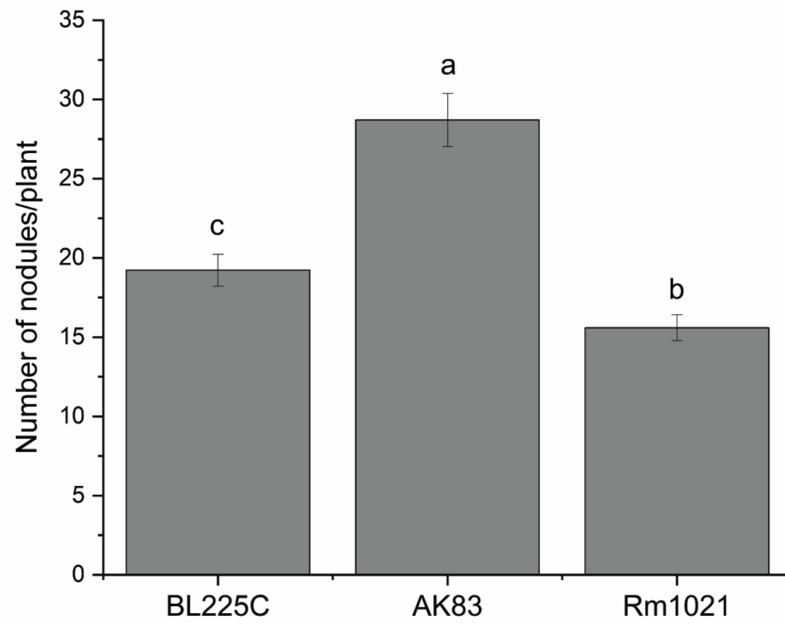

D
nmoles ethylene/h/plant
250
200
150
100
50
0
abcd
abc
d
abc
bcd
a
abc
ab
cd
abc
abc
d
abc
d
KH46
CCMM B554
T073
Rm41
HM006
GR4
2011
USDA1157
KH35c
SM11
RU11/001
M270
AK58
C-

Figure S3. Number of nodules in the three competitions. A) Total nodules per plant. B) Total mixed nodules per plant. Different letters indicate significant differences between treatments ($p < 0.05$).



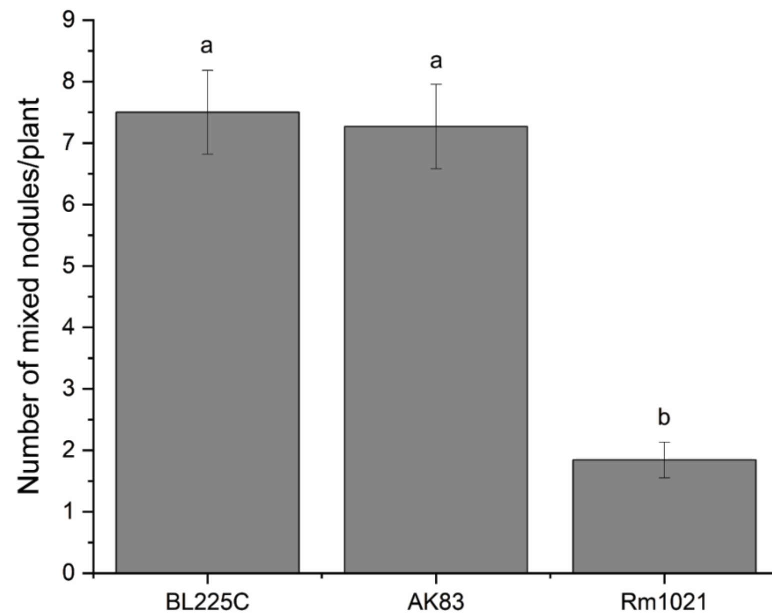
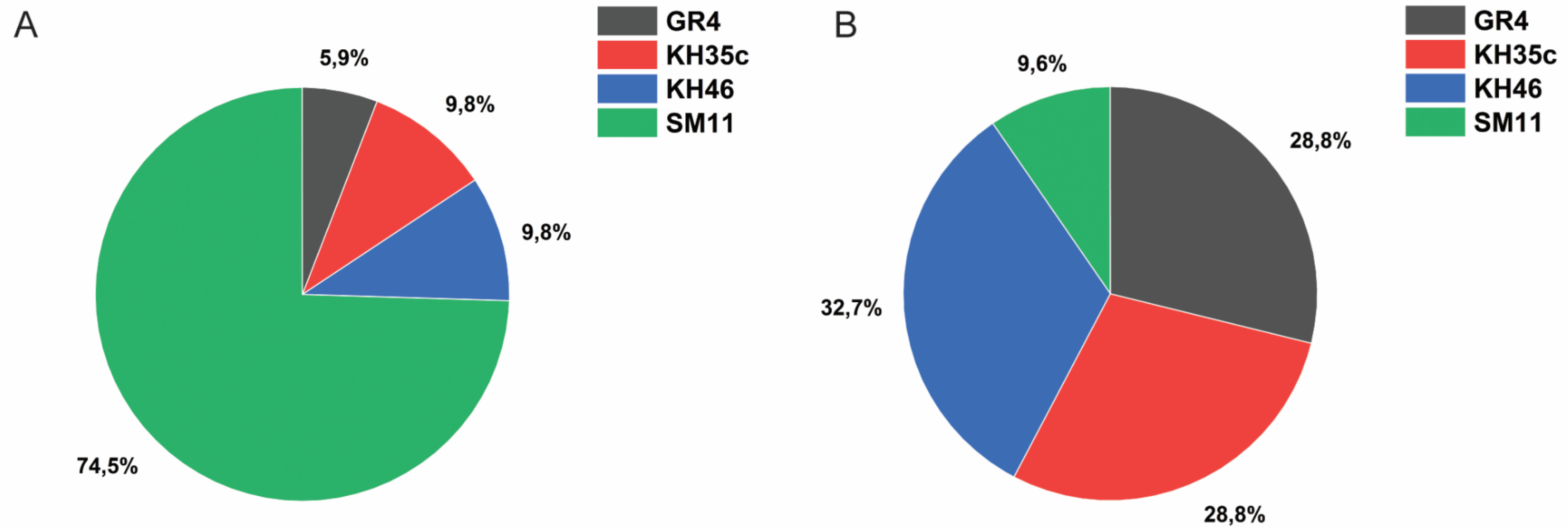


Figure S4. Distribution of 51 best k-mers tagged-CDSs in *S. meliloti* GR4, KH35c, KH46 and SM11. A) Distribution of unannotated CDSs among four *S. meliloti* strains and distribution of unannotated CDSs among replicons of *S. meliloti* strains C) GR4, D) KH35c, E) KH46 and F) SM11. B) Distribution of orthologous genes hits among four strains *S. meliloti* strains and distribution of orthologous genes among replicons of *S. meliloti* strains G) GR4, H) KH35c, I) KH46 and J) SM11.



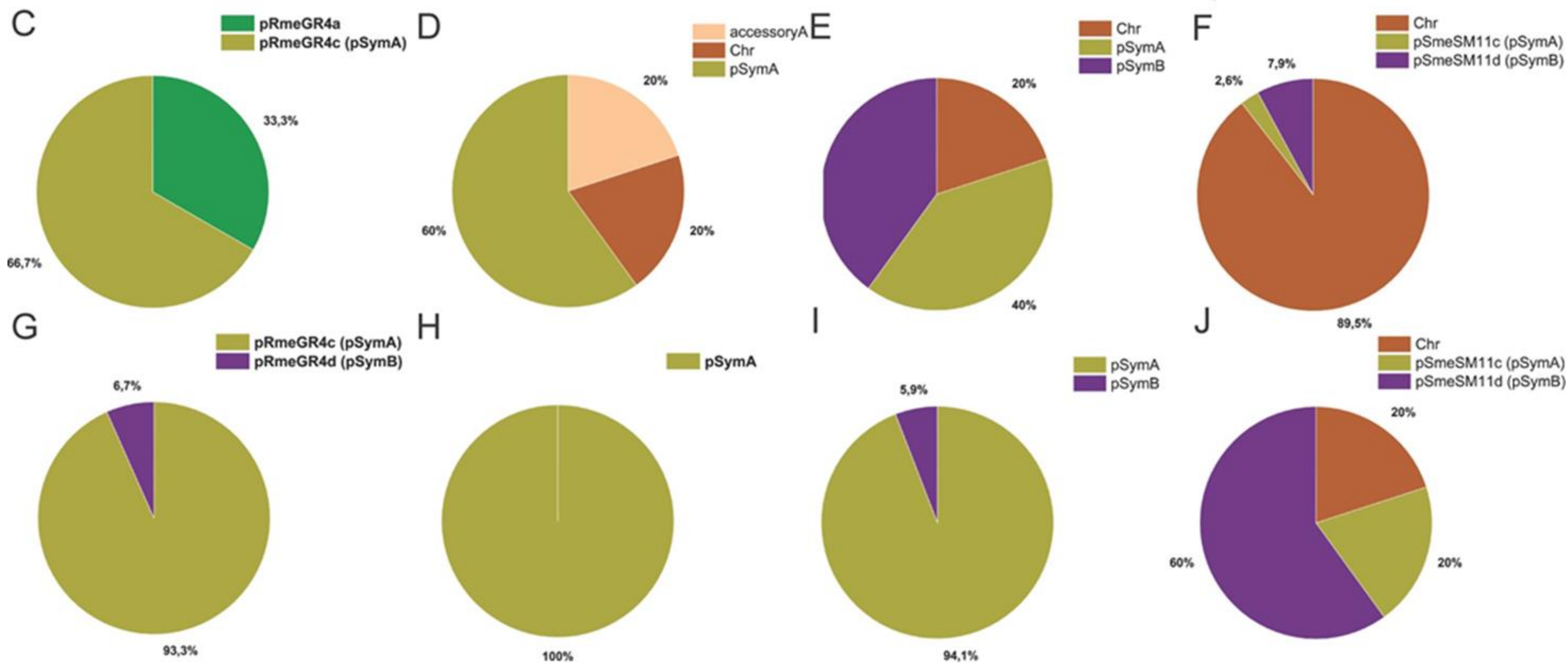
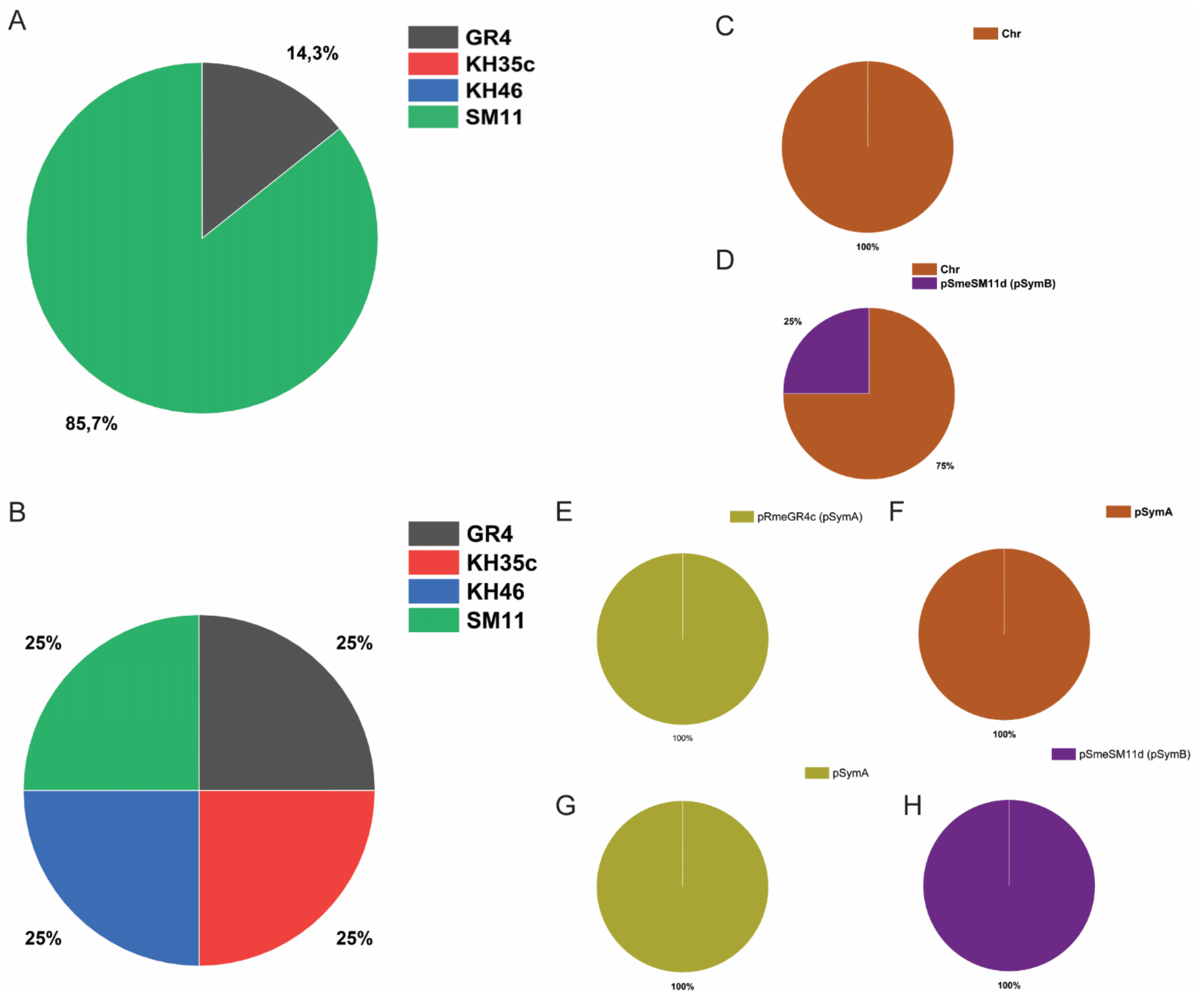

C
pRmeGR4a
pRmeGR4c (pSymA)
33,3%
66,7%
D
accessoryA
Chr
pSymA
20%
20%
60%
E
Chr
pSymA
pSymB
20%
40%
F
Chr
pSmeSM11c (pSymA)
pSmeSM11d (pSymB)
2,6%
7,9%
89,5%
G
pRmeGR4c (pSymA)
pRmeGR4d (pSymB)
6,7%
93,3%
H
pSymA
100%
I
pSymA
pSymB
5,9%
94,1%
J
Chr
pSmeSM11c (pSymA)
pSmeSM11d (pSymB)
20%
20%
60%

Figure S5. Distribution of 10 k-mers tagged-putative regulatory regions in *S. meliloti* GR4, KH35c, KH46 and SM11. A) Distribution of regulatory region hits of unannotated CDSs among four *S. meliloti* strains and among replicons of *S. meliloti* strains C) GR4 and D) SM11. B) Distribution of putative regulatory region hits of orthologous gene hits among four *S. meliloti* strains and among replicons of *S. meliloti* strains E) GR4, F) KH35c, G) KH46 and H) SM11.



Supplementary tables

Table S1. ***Sinorhizobium meliloti*** **strains used in this work.**

Strains	Source/Description	Genbank assembly codes	Reference
AK83	Geographic location: Kazakhstan; Host: *Medicago falcata*	GCA_000147795.3 (11)	(12)
1021	SU47 *str*-21	GCA_000006965.1 (13)	(14)
BL225C	Geographic location: Italy; Host: *Medicago sativa*	GCA_000147775.3 (11)	(15)
KH46	Geographic location: France; Host: *Medicago truncatula*	GCF_002197465.1 (16, 17)	(16)
CCMM B554	Geographic location: Morocco; Host: *Medicago arborea*	GCA_002215195.1 (18)	(19)
T073	Geographic location: Tunisia; Host: *Medicago truncatula*	GCA_002197145.1 (16, 17)	(16)
Rm41	Geographic location: Hungary; Host: *Melilotus/Medicago*	GCA_000304415.1 (20)	(21)
HM006	Geographic location: France; Host: *Medicago truncatula*	GCA_002197165.1 (16, 17)	(16)
GR4	Geographic location: Spain; Host: agricultural field	GCA_000320385.2 (22)	(22)
2011	SU47	GCA_000346065.1 (23)	-
USDA1157	Geographic location: USA, California; Host: *Medicago sativa*	GCF_002197025.1 (17)	(17)
KH35c	Geographic location: France; Host: *Medicago truncatula*	GCA_002197105.1 (16, 17)	(16)
SM11	Geographic location: Germany; Host: agricultural field	GCA_000218265.1 (24)	(25)

RU11/001	Geographic location: Germany; Host: *Medicago sativa*	GCA_001050915.2 (26)	(27)
M270	Geographic location: Jordan; Host: *Medicago truncatula*	GCA_002197085.1 (16)	(16)
AK58	Geographic location: Kazakhstan; Host: *Medicago falcata*	GCA_000473425.1 (28)	(12)

Table S2. Single nodule occupancy of *S. meliloti* tested strains in competition experiments versus *S. meliloti* strains Rm1021, AK83 and BL225C. Different letters indicate statistically significant differences (Kuskal-Wallis and Dunn test, $p<0.05$) within a competition assay (columns; vs BL225C, vs AK83, vs Rm1021).

Strains	vs BL225C	vs AK83	vs Rm1021
AK58	41.2% ^{abcd}	45.7% ^{a}	82.7 % ^{cd}
CCMM B554	42.9% ^{abc}	28.6% ^{ab}	65.0 % ^{acd}
GR4	66.7% ^{a}	63.9% ^{a}	93.4 % ^{d}
HM006	46.4% ^{ab}	25.2% ^{ab}	58.6 % ^{abcd}
KH35c	68.3% ^{a}	28.4% ^{abc}	89.3 % ^{d}
KH46	68.9% ^{a}	7.0% ^{bc}	90.0 % ^{d}
M270	37.5% ^{abcd}	6.8% ^{bc}	37.0 % ^{abc}
2011	15.5% ^{bcd}	8.6% ^{bc}	84.8 % ^{d}
Rm41	13.9% ^{bcd}	1.7% ^{c}	50.5 % ^{abcd}
RU11/001	39.4% ^{abcd}	30.0% ^{ab}	89.0 % ^{d}
SM11	63.4% ^{a}	29.1% ^{ab}	86.6 % ^{d}
T073	0.4% ^{d}	1.8% ^{c}	0.00 % ^{b}
USDA 1157	8.3% ^{cd}	9.1% ^{bc}	19.7 % ^{ab}

Table S3. Linear regression models for the three competition experiments, performed by PhenotypeSeeker with 3-fold train/test splits of samples. The averaged model evaluation metrics of both training and test set are reported.

Dataset		The mean squared error	The coefficient of determination (R^2)	The Pearson correlation and p-value	The Spearman correlation coefficient and p-value	Total K-mers (p-value < 0.05)	Range p-value
Vs Rm1021	Training set	0.02	0.74	0.86, 0.0	0.98, 0.0	439886	4.99E-0^2 – 4.31E-0^3
	Test set	0.01	0.84	0.96, 0.04	1.0, 0.0		
Vs BL225C	Training set	0.0	0.98	0.99, 0.0	0.97, 0.0	182804	4.98E-0^2 – 1.35E-0^5
	Test set	0.0	0.96	0.99, 0.01	0.89, 0.11		
Vs AK83	Training set	0.0	0.99	0.99, 0.0	0.92, 0.0	292884	4.97E-0^2 – 1.05E-0^3
	Test set	0.04	0.11	0.8, 0.2	0.77, 0.23		

Table S4. List of top k-mers (raw data k-mers). (file excel available on: https://doi.org/10.1101/2020.09.15.298034)

Table S5. Genes hits identified by 51 best k-mers (raw data kmers). (file excel available on: https://doi.org/10.1101/2020.09.15.298034)

Table S6. List of COGs codes.

COG ID	COG name
J	Translation, ribosomal structure and biogenesis
A	RNA processing and modification
K	Transcription
L	Replication, recombination and repair
B	Chromatin structure and dynamics
D	Cell cycle control, cell division, chromosome partitioning
Y	Nuclear structure
V	Defense mechanisms
T	Signal transduction mechanisms
M	Cell wall/membrane/envelope biogenesis
N	Cell motility
Z	Cytoskeleton
W	Extracellular structures
U	Intracellular trafficking, secretion, and vesicular transport
O	Post-translational modification, protein turnover, chaperones
X	Mobilome: prophages, transposons
C	Energy production and conversion
G	Carbohydrate transport and metabolism
E	Amino acid transport and metabolism
F	Nucleotide transport and metabolism
H	Coenzyme transport and metabolism
I	Lipid transport and metabolism
P	Inorganic ion transport and metabolism

Q	Secondary metabolism biosynthesis, transport and catabolism
R	General function prediction only
S	Function unknown

Table S7. Regulatory region hits identified by 10 best k-mers. (file excel available on: https://doi.org/10.1101/2020.09.15.298034)

Table S8. Strains and plasmids used in this work

Species	Strains (or plasmids)	Source/Description	Resistances	Reference
Sinorhizobium meliloti	AK83			(12)
	1021	SU47 *str*-21	Str[1]	(14)
	BL225C			(15)
	KH46			(16)
	CCMM B554			(19)
	T073			(16)
	Rm41			(21)
	HM006			(16)
	GR4			(22)
	2011	SU47	Str	-
	USDA1157			(17)
	KH35c			(16)
	SM11			(25)
	RU11/001		Str	(27)
	M270			(16)
	AK58			(12)
	BM685	AK83 pBHR - mRFP	Rif[2] & Tc[3]	(29)
	BM687	1021 pBHR - mRFP	Str & Tc	(29)
	GE0346	BL225C pBHR – mRFP	Rif & Tc	This work

	GE0323	KH46 pHC60	Rif & Tc	This work
	GE0326	CCMM B554 pHC60	Rif & Tc	This work
	GE0327	T073 pHC60	Rif & Tc	This work
	GE0328	Rm41 pHC60	Rif & Tc	This work
	GE0329	HM006 pHC60	Rif & Tc	This work
	GE0330	GR4 pHC60	Rif & Tc	This work
	GE0339	2011 pHC60	Str & Tc	This work
	GE0341	USDA1157 pHC60	Rif & Tc	This work
	GE0342	KH35c pHC60	Rif & Tc	This work
	GE0345	SM11 pHC60	Rif & Tc	This work
	GE0357	RU11/001 pHC60	Rif & Tc	This work
	GE0359	M270 pHC60	Rif & Tc	This work
	GE0360	AK58 pHC60	Rif & Tc	This work
Escherichia coli	BM266	S17-1 λpir pHC60	Tc	(29)
	BM679	S17-1 λpir pBHR- mRFP	Tc	(29)
Plasmids	pBHR-mRFP	Constitutive expression of RFP	Tc	(2)
	pHC60	Constitutive expression of GFP	Tc	(1)

Supplemental material references

1. Cheng HP, Walker GC. 1998. Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. J Bacteriol 180:5183–5191.

2. Smit P, Raedts J, Portyanko V, Debellé F, Gough C, Bisseling T, Geurts R. 2005. NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. Science (80-) 308:1789–1791.

3. Pini F, Spini G, Galardini M, Bazzicalupo M, Benedetti A, Chiancianesi M, Florio A, Lagomarsino A, Migliore M, Mocali S, Mengoni A. 2014. Molecular phylogeny of the nickel-resistance gene nreB and functional role in the nickel sensitive symbiotic nitrogen fixing bacterium *Sinorhizobium meliloti*. Plant Soil 377:189–201.

4. Ramachandran VK, East AK, Karunakaran R, Downie JA, Poole PS. 2011. Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biol 12:R106.

5. R Development Core Team. 2011. R: A language and environment for statistical computing. R Found Stat Comput.

6. Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

7. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.

8. Tamura K, Nei M, Kumar S. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A 101:11030–11035.

9. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549.

10. Pagès H, Aboyoun P, Gentleman R, DebRoy S. 2017. Biostrings: Efficient manipulation of biological strings. R Packag version 2460.

11. Galardini M, Mengoni A, Brilli M, Pini F, Fioravanti A, Lucas S, Lapidus A, Cheng J, Goodwin L, Pitluck S, Land M, Hauser L, Woike T, Mikhailova N, Ivanova N, Daligault

H, Bruce D, Detter C, Tapia R, Han C, Teshima H, Mocali S, Bazzicalupo M, Biondi EG. 2011. Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. BMC Genomics 12:235.

12. Giuntini E, Mengoni A, De Filippo C, Cavalieri D, Aubin-Horth N, Landry CR, Becker A, Bazzicalupo M. 2005. Large-scale genetic variation of the symbiosis-required megaplasmid pSymA revealed by comparative genomic analysis of *Sinorhizobium meliloti* natural strains. BMC Genomics2005/11/15. 6:158.

13. Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, Bothe G, Boutry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Dreano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernandez-Lucas I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetelle D, Purnelle B, Ramsperger U, Surzycki R, Thebault P, Vandenbol M, Vorholter FJ, Weidner S, Wells DH, Wong K, Yeh KC, Batut J. 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science (80-) 293:668–672.

14. Meade HM, Long SR, Ruvkun GB, Brown SE, Ausubel FM. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. J Bacteriol1982/01/01. 149:114–122.

15. Carelli M, Gnocchi S, Fancelli S, Mengoni A, Paffetti D, Scotti C, Bazzicalupo M. 2000. Genetic diversity and dynamics of *Sinorhizobium meliloti* populations nodulating different alfalfa cultivars in Italian soils. Appl Environ Microbiol 66:4785–4789.

16. Sugawara M, Epstein B, Badgley BD, Unno T, Xu L, Reese J, Gyaneshwar P, Denny R, Mudge J, Bharti AK, Farmer AD, May GD, Woodward JE, Medigue C, Vallenet D, Lajus A, Rouy Z, Martinez-Vaz B, Tiffin P, Young ND, Sadowsky MJ. 2013. Comparative genomics of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. Genome Biol2013/02/22. 14:R17.

17. Nelson M, Guhlin J, Epstein B, Tiffin P, Sadowsky MJ. 2018. The complete replicons of 16 *Ensifer meliloti* strains offer insights into intra- and inter-replicon gene transfer, transposon-associated loci, and repeat elements. Microb Genom 4:e000174.

18. Nagymihály M, Vásarhelyi BM, Barrière Q, Chong T-M, Bálint B, Bihari P, Hong K-W, Horváth B, Ibijbijen J, Amar M, Farkas A, Kondorosi É, Chan K-G, Gruber V, Ratet P, Mergaert P, Kereszt A. 2017. The complete genome sequence of *Ensifer meliloti* strain CCMM B554 (FSM-MA), a highly effective nitrogen-fixing microsymbiont of *Medicago truncatula* Gaertn. Stand Genomic Sci 12:75.

19. Kazmierczak T, Nagymihály M, Lamouche F, Barrière Q, Guefrachi I, Alunni B, Ouadghiri M, Ibijbijen J, Kondorosi É, Mergaert P, Gruber V. 2017. Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium* strains. Mol Plant Microbe Interact 30:399–409.

20. Weidner S, Baumgarth B, Göttfert M, Jaenicke S, Pühler A, Schneiker-Bekel S, Serrania J, Szczepanowski R, Becker A. 2013. Genome Sequence of *Sinorhizobium meliloti* Rm41. Genome Announc 1:e00013-12.

21. Szende K, Ördögh F. 1960. Die Lysogenie von *Rhizobium meliloti*. Naturwissenschaften 47:404–405.

22. Martinez-Abarca F, Martinez-Rodriguez L, Lopez-Contreras JA, Jimenez-Zurdo JI, Toro N. 2013. Complete genome sequence of the alfalfa symbiont *Sinorhizobium*/*Ensifer meliloti* strain GR4. Genome Announc 1:e00174-12.

23. Sallet E, Roux B, Sauviac L, Jardinaud MF, Carrere S, Faraut T, de Carvalho-Niebel F, Gouzy J, Gamas P, Capela D, Bruand C, Schiex T. 2013. Next-generation annotation of prokaryotic genomes with EuGene-P: application to *Sinorhizobium meliloti* 2011. DNA Res 20:339–354.

24. Schneiker-Bekel S, Wibberg D, Bekel T, Blom J, Linke B, Neuweger H, Stiens M, Vorholter FJ, Weidner S, Goesmann A, Puhler A, Schluter A. 2011. The complete genome sequence of the dominant *Sinorhizobium meliloti* field isolate SM11 extends the *S. meliloti* pan-genome. J Biotechnol 155:20–33.

25. Stiens M, Schneiker S, Keller M, Kuhn S, Pühler A, Schlüter A. 2006. Sequence analysis of the 144-kilobase accessory plasmid pSmeSM11a, isolated from a dominant *Sinorhizobium meliloti* strain identified during a long- term field release experiment. Appl Environ Microbiol 72:3662–3672.

26. Wibberg D, Blom J, Ruckert C, Winkler A, Albersmeier A, Puhler A, Schluter A, Scharf BE. 2013. Draft genome sequence of *Sinorhizobium meliloti* RU11/001, a model organism for flagellum structure, motility and chemotaxis. J Biotechnol 168:731–733.

27. Krupski G, Götz R, Ober K, Pleier E, Schmitt R. 1985. Structure of complex flagellar filaments in *Rhizobium meliloti*. J Bacteriol 162:361–366.

28. Galardini M, Bazzicalupo M, Biondi E, Brambilla E, Brilli M. 2013. Permanent draft genome sequences of the symbiotic nitrogen fixing *Ensifer meliloti* strains BO21CC and AK58. Stand Genomic Sci 9:325–333.

29. Checcucci A, Azzarello E, Bazzicalupo M, Galardini M, Lagomarsino A, Mancuso S, Marti L, Marzano MC, Mocali S, Squartini A, Zanardo M, Mengoni A. 2016. Mixed nodule infection in *Sinorhizobium meliloti–Medicago sativa* symbiosis suggest the presence of cheating behavior. Front Plant Sci 7:835.

Chapter 4

Deciphering the Symbiotic Plant Microbiome: Translating the Most Recent Discoveries on Rhizobia for the Improvement of Agricultural Practices in Metal-Contaminated and High Saline Lands

Agnese Bellabarba[1], Camilla Fagorzi[2], George C. diCenzo[3], Francesco Pini[1], Carlo Viti[1] and Alice Checcucci [2,4]

[1] Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, 50144 Florence, Italy

[2] Department of Biology, University of Florence, 50019 Sesto Fiorentino, Italy

[3] Department of Biology, Queen's University, Kingston, ON K7L 3N6, Canada

[4] Department of Agricultural and Food Science (DISTAL), University of Bologna, 40127 Bologna, Italy

Received on Agronomy 1 July 2019; Accepted: 5 September 2019; Published: 10 September 2019

Personal contribution: original draft preparation, literature review sections

Agronomy 2019, 9, 529; doi:10.3390/agronomy9090529

agronomy

MDPI

Review

Deciphering the Symbiotic Plant Microbiome: Translating the Most Recent Discoveries on Rhizobia for the Improvement of Agricultural Practices in Metal-Contaminated and High Saline Lands

Agnese Bellabarba[1], Camilla Fagorzi[2], George C. diCenzo[3], Francesco Pini[1], Carlo Viti[1] and Alice Checcucci [2,4]

[1] Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, 50144 Florence, Italy

[2] Department of Biology, University of Florence, 50019 Sesto Fiorentino, Italy

[3] Department of Biology, Queen's University, Kingston, ON K7L 3N6, Canada

[4] Department of Agricultural and Food Science (DISTAL), University of Bologna, 40127 Bologna, Italy

* Correspondence: alice.checcucci2@unibo.it

Abstract: Rhizosphere and plant-associated microorganisms have been intensely studied for their beneficial effects on plant growth and health. These mainly include nitrogen-fixing bacteria (NFB) and plant-growth promoting rhizobacteria (PGPR). This beneficial fraction is involved in major functions such as plant nutrition and plant resistance to biotic and abiotic stresses, which include water deficiency and heavy-metal contamination. Consequently, crop yield emerges as the net result of the interactions between the plant genome and its associated microbiome. Here, we provide a review covering recent studies on PGP rhizobia as effective inoculants for agricultural practices in harsh soil, and we propose models for inoculant combinations and genomic manipulation strategies to improve crop yield.

Keywords: soil bioremediation; high-salinity soil; plant beneficial microbes; rhizobia; microbial inoculants; plant-growth promoting rhizobacteria (PGPR)

1. Introduction

The last ten years have witnessed a number of discoveries and an increased awareness of the importance of the microbiome for the health and the growth of host macroorganisms [1,2]. Plants and their related microbiota can be considered holobionts, complex systems ruled by interdependent and composite interactions [3–5]. Indeed, plants are colonized by an astounding number of microorganisms that can reach numbers much greater than those of plant cells. This is especially relevant for the rhizosphere, the thin layer of soil surrounding and influenced by plant roots, where it is possible to observe a staggering diversity of microorganisms; a single gram of rhizospheric soil hosts tens of thousands of distinct microbial species [6,7]. Plants influence the composition of their rhizosphere microbiota through the production of root exudates [8], which differ in space and time [9], contributing to the positive selection of plant-growth promoting (PGP) and beneficial bacteria [10]. Indeed, it has been suggested that plants may have evolved the beneficial trait of secreting specific compounds to recruit protective microorganisms in response to pathogen attacks [11,12].

Plant-based bacterial selection relies on the microorganism already present in the soils where they are grown. Therefore, crop productivity could be increased by modifying root microbiota with microbial inoculants, which may be composed of a single strain or a consortia of different PGP rhizobacteria (PGPR) [13,14].

Within the beneficial plant microbiota, rhizobia constitute one of the most studied fractions [15]. The “rhizobium” definition is based on the ability to induce the formation of root/stem nodules in leguminous plants [16]. However, rhizobia are found not only on legumes, but are also found in association with several plant species [17]. Within legume nodules, rhizobia differentiate into bacteroids and synthesize a protein complex called nitrogenase that converts atmospheric dinitrogen to ammonia (Biological Nitrogen Fixation, BNF) [18]. The produced ammonia is then transferred to the host plant to sustain its biosynthetic pathways [16]. The establishment of the symbiotic interaction between the nitrogen fixing rhizobia and leguminous plants is highly regulated and begins with mutual recognition between the plant and the rhizobia present in the rhizosphere. To date, rhizobia have been identified in two bacterial classes, the Alphaproteobacteria and the Betaproteobacteria. In the Alphaproteobacteria, rhizobial strains are present in the genera *Sinorhizobium* (syn. *Ensifer*), *Rhizobium*, *Mesorhizobium*,

Bradyrhizobium, *Azorhizobium*, *Methylobacterium*, *Devosia*, *Ochrobactrum*, *Aminobacter*, *Microvirga*, *Shinella*, and *Phyllobacterium*. In the Betaproteobacteria, rhizobia are present within strains of the genera *Paraburkholderia*, *Cupriavidus*, and *Trinickia* [19–21]. While the beta-rhizobia are mainly found in association with tropical legumes, the alpha-rhizobia appear to be more widespread and nodulate tropical to temperate legumes including pasture, tree, and grain legumes. The alpha-rhizobia have received more research attention than the beta-rhizobia; among the alpha-rhizobia, *Sinorhizobium* (syn. *Ensifer*) is likely the most studied genus, followed by the genera *Rhizobium* and *Bradyrhizobium* [22,23].

In this review, we highlight the most recent studies on PGP rhizobia isolated from, and adapted to, drought-affected and metal-contaminated soils and their possible use as effective inoculants for legumes grown in harsh agricultural soils. We cover the identification of the genetic determinants of their tolerance, as well as the mechanisms that allow rhizobia to survive and to improve host plant growth in harsh soils. Models of inoculant combinations and genomic manipulation strategies for the improvement of crop yield are discussed.

2. The Need for Rhizobial Inoculants

The demand for plant proteins for human nutrition has increased tremendously over the last fifteen years. This can be related to: (i) Demographic growth and urbanization, (ii) the limited land areas that can be used for the production of food crops while farming systems are switching to specialized but unsustainable cereal production (for market competitiveness), and (iii) decreases in animal protein production due to shortage of irrigation and/or rainfall water especially in arid areas. The demand for plant proteins can be met in part through the cultivation of protein-rich leguminous crops. Additionally, legumes can help improve soil fertility through symbiotic nitrogen fixation, and they can help protect ground water from toxicity resulting from excessive application of N-fertilizers [24].

In the past three decades, eco-sustainable agronomic practices have been employed in an attempt to replace chemical fertilizers and pesticide-based agriculture [25,26]. Therefore, the exploitation of beneficial microorganisms as biofertilizer has become of primary importance [27]. In particular, rhizobial bioformulations could partially or completely substitute mineral nitrogen fertilizers [28,29]. *Rhizobium*-legume symbioses provide more than half of the world's biologically fixed nitrogen [30], and it was reported that rhizobial nitrogen fixation introduces 40–48 million tonnes of nitrogen into agricultural systems each year [31]. The impact of BNF on the global agricultural economy was estimated to be worth the equivalent of USD160–180

billion [32]. *Rhizobium* inoculants are already widely used in agriculture, providing one of the most cost-effective ways to boost legume performances [33,34]. However, with a few exceptions, the last fifteen years has seen only small enhancements in the production of traditionally grown grain legumes such as fava bean, chickpea, lentils, or common beans [35]. Generally, yield instability is the main constraint for increasing plant productivity. Thus, special attention must be given to the factors that reduce soil quality and decrease plant yield.

Recent works have highlighted that microbial species associated with plants (rhizobial and non-rhizobial strains, including mycorrhizal fungi) can positively influence plant tolerance to water deficiency [36]. This is due to their PGP features such as indoleacetic acid, siderophore production, phosphate and zinc solubilization [37], and the synthesis of 1-aminocyclopropane-1-carboxylate (ACC) deaminase [38], which are more evident and easily identified in stressful conditions [39]. Nevertheless, legumes are strongly affected by water deficit. In particular, BNF appears to be more sensitive to water deficit than other physiological functions such as photosynthesis or nutrient uptake [40]. Sometimes, this results in impaired nodule development [41] or the accumulation of small, generally organic, osmolytes called compatible solutes [42].

Aside from water deficiency and soil nutrient depletion, heavy-metal contamination due to anthropic activities (agricultural and industrial practices) or the weathering of metal-enriched rocks have recently increased exponentially, becoming a worldwide problem for crop productivity [43–45]. Generally, plant-associated microbes can contribute to a plant's ability to perform phytoextraction (accumulation of toxic compounds in the plant tissues) and phytostabilization (adsorption through the roots and conversion into harmless compounds). In legumes, phytostabilization is the key process when considering the phytoremediation of contaminated soils [46,47], and their associated rhizobia can promote chemical transformation and the chelation of heavy-metal compounds [48] (Figure 1). Therefore, plant growth and agricultural yield is related not only to the plant genotype and the soil condition, but, especially for legumes, their associated microbiota also play important roles [49,50]. As such, the selection of rhizobial strains resistant to water deficiency and capable of alleviating metal phytotoxicity could be a crucial strategy to improve the yield of legumes growth in arid or in metal-contaminated soils.

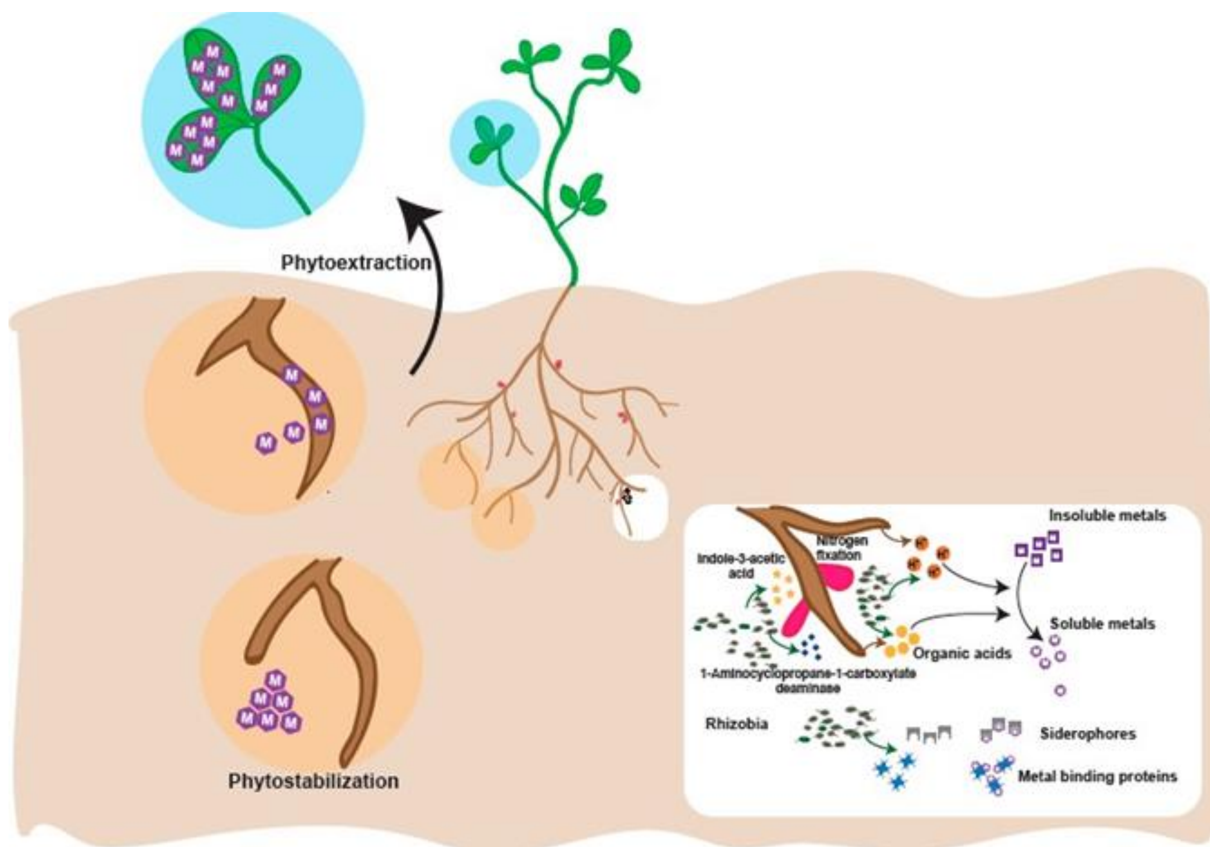


Figure 1. The mechanisms involved in bioremediation of heavy-metal contaminated soil and the contribution of PGP rhizobia

3. Plant Growth Promoting Rhizobia in Saline and Harsh Soil

Salinity due to water deficiency is one of the largest environmental constraints for plant growth and productivity in stricken regions [51,52]. It was estimated that almost 40% of the world's lands can potentially become arid or semi-arid [52], most of which are located in the tropics and in the Mediterranean area [53,54]. The persistent increase of anthropic activities (such as poor agricultural practices) decreases soil water availability, alters the soil microbiota, and reduces the nutritive value of soils [55]. The progressive salinization of soil may cause several stresses to the plants that decrease their growth. This negatively affects crop productivity, hindering the agricultural economy of developing countries [56].

Desiccated soils can lead to alterations in plant homeostasis due to a reduction in the osmotic potential and inappropriate ionic distribution [56]. The alteration in nutrient supply and the resulting nutritional imbalance induces a loss of turgor pressure and the growth of smaller

leaves [56]. Furthermore, increasing Na+ and Cl− concentrations commonly leads to the formation of burn-like lesions that alter leaf transpiration [57]. Continued growth interruption results in less efficient photosynthesis, respiratory variations, premature senescence, and the loss of cellular integrity leading to plant death [58,59]. Salt tolerance among legume species can fluctuate and is dependent on the chemical features of the soil, the climatic conditions, and the growth stage of the plant [53,60,61]. Tree legumes such as *Prosopis* [53,62] and *Acacia* spp. [53,63] are highly tolerant to salinity, as are the grain legumes *Glycine max* [51] and *Vicia faba* [61,64]. On the contrary, *Cicer arietinum*, *Phaseolus vulgaris*, and *Pisum sativum* are known to be extremely sensitive to salt stress [64–67]. BNF can also be strongly affected by a lack of water, influencing the mutual symbiotic interaction between the host plants and their associated rhizobia [53]. Generally, salt and osmotic stresses lead to a decrease in rhizobium nodulation, affecting the early stages of the symbiotic process [68–71]. This commonly includes poor root colonization by the bacteria and reduced curling and deformation of roots hairs [53,72,73]. Furthermore, a decrease in nitrogenase activity under drought/salinity stress is commonly attributed to a reduction in nodule respiration [67,74–76], which reduces the synthesis of cytosolic proteins such as leghemoglobin [67,76,77].

It was reported that host legumes are less tolerant to salt than their associated rhizobia [71,73,78]. For example, *Rhizobium leguminosarum* bv. trifolii TA-1 can tolerate up to 350 mM of NaCl in vitro [78]. The highest levels of salt tolerance seem to be associated with rhizobia isolated from woody legumes (e.g., *Prosopis*, *Acacia*, and *Leucaena*), which are tolerant of up to 850 mM of NaCl [73,79,80]. For example, Sakrouhi and collaborators [81] isolated 20 symbiotic N_2-fixing bacteria from *Acacia tortilis* and *Acacia gummifera* that were able to grow in high salt media (400 mM). The intracellular accumulation of low-molecular-weight organic solutes, osmo-protectants, is an osmotic adaptation mechanism used by a large variety of bacterial species [53]. These compounds are acquired through de novo synthesis or uptake from the environment, and they can be accumulated to high intracellular concentrations without interfering with cellular processes [56], such as in the *S. meliloti* 102F34 strain [82–84]. The accumulation of poly-hydroxybutyrate (PHB) has also been reported as a protective measure to help rhizobia survive in high saline environments [85]. Following a decrease in osmolarity, the osmo-protectants are released by bacteria into the surrounding environment and actively recovered by plants, which are unable to synthesize them de novo [84]. The successful uptake of these compounds by the plants improve their growth under osmotic stress [56]. Additionally, intracellular trehalose accumulation by *R. leguminosuarum* seems to be involved in metabolic osmoregulation of the host plants [86,87]. The intracellular accumulation of

glycine betaine was identified as one of the most frequent osmotic stress responses of rhizobia [88,89]. Several lines of evidence suggest that this process plays a role in maintaining bacteroid nitrogenase activity in *Medicago sativa* nodules [53,90–92].

Nowadays, the massive use of fertilizers to offset the effect of soil nutrient loss (also due to salinization) on crop productivity seems to be the preferred solution. This choice has progressively contributed to the deterioration of the soils that were already compromised by intense agricultural practices. An alternative is to make use of bacterial inoculants adapted to these harsh conditions. Studying the bacterial communities associated with plants growing in saline soils, and the underlying mechanism of their effectiveness, can be a good starting point for the use of microbial inoculant in agricultural practices to reduce saline stress [56,93]. To begin addressing this, the genomes of several rhizobia that nodulate plants in harsh environments were sequenced to identify stress-adaption genes. Examples include: *Rhizobium* sp. LCM 4573 (a salt-tolerant rhizobium from Senegalese soil [94]), *R. leucaenae* (a stress-tolerant species nodulating plants in tropical acid soils [95]), *S. meliloti* AK21 (from the Aral Sea Region that experiences saline and drought conditions [96]) and *Ensifer aridi* (isolated from arid soils of diverse deserts [97]). Transcriptomic and proteomic studies have been instrumental in highlighting how rhizobia respond to environmental stresses (a detailed list is provided elsewhere [98]). Studies with *Mesorhizobium loti* suggested heat shock results in a global downregulation of protein expression possibly to conserve energy [99], while salt stress leads to over-expression of ABC transporters and genes associated with nucleotide transport and metabolism [100]. Similarly, characterization of the acid stress response in *S. meliloti* suggested that this stress results in an elevated metabolic respiration rate [101]. However, to date only a few studies have examined the responses of differentiated bacteroids to environmental stress [102,103], although several studies have examined the responses of the plant partner [104]. Additional studies characterizing the response of bacteroids to environmental stresses would be beneficial to complement the free-living datasets.

The agricultural significance of uncovering the genetic and metabolic basis of stress resistance in rhizobia and other PGP bacteria has been emphasized from at least the early 1990s [105]. In fact, numerous studies have demonstrated that genetically modifying rhizobia can increase or decrease their symbiotic abilities in stressful environments [106]. The grain yield of common bean plants grown in drought conditions was significantly higher when inoculated with a *Rhizobium etli* strain overexpressing otsA, encoding a trehalose-6-phosphate [107]. Similar results were obtained for maize plants inoculated with a non-rhizobium diazotroph

Azospirillum brasilense strain overexpressing a trehalose biosynthesis gene [108]. In contrast, soybean plants inoculated with a *Bradyrhizobium japonicum* putA mutant that is unable to catabolize proline produced fewer seeds than plants inoculated with the wild-type parental strain when grown in moderate drought conditions [109]. *S. meliloti* strains overexpressing betS displayed improved nitrogen fixation phenotypes during salt stress [110], while salt-sensitive *Rhizobium trop*ici mutants were poor symbionts even in the absence of stress [111]. Finally, ACC-deaminases encoded by some rhizobia can reduce the overproduction of the plant gas hormone ethylene during abiotic stresses [38,112], reducing the deleterious effect of ethylene and thus improving plant growth [113].

The inoculation of plants with microbial communities has also been shown to improve plant tolerance to environmental stresses. The co-inoculation of soybean (*Glycine max*) with *Chryseobacterium balustinum* Aur9 and *Ensifer* (*Sinorhizobium*) *fredii* SMH12 led to increased symbiotic performance under saline conditions (25 mM NaCl) [114]. In the same study, co-inoculation of common bean (Phaseolus vulgaris L.) with *R. tropici* CIAT899 and *C. balustinum* Aur9 enhanced bean growth in both saline (25 mM NaCl) and control conditions compared to single strain inoculation [114]. Moreover, co-inoculation of *Rhizobium phaseoli* M6 and M9, *Pseudomonas syringae* Mk1, *Pseudomonas fluorescens* Mk20, and *Pseudomonas fluorescens* biotype G Mk25 strains decreased the effects of salinity stress in bean, enhancing its nodulation process in vitro and in fields conditions [115,116].

4. Plant Growth Promoting Rhizobia in Heavy Metal Contaminated Soil

Anthropogenic activities, such as the use of fertilizers and pesticides in agricultural soils, the production of sewage sludge waste, and industrial and mining activities, are responsible for the accumulation of toxic heavy metals in the food chain [117]. Low concentrations of metals such as zinc (Zn), copper (Cu), iron (Fe), nickel (Ni), manganese (Mn), molybdenum (Mo) and cobalt (Co) are necessary for the metabolism of all organisms [118]. However, high concentrations of these metals, as well as the long-term persistence in the soil of elements such as cadmium (Cd), lead (Pb), and arsenic (As), negatively affect the composition of microbial communities [119], the dynamics of the rhizosphere niche [120], and the growth, the biomass, and the photosynthesis of plants [121]. Plant species used for the remediation of heavy metal polluted sites represent an environment-friendly, aesthetically appealing, and cost-effective solution. Legumes may be ideal species for bioremediation as surveys of plant species surviving

in long-term metal-contaminated environments have shown legumes to account for a dominant portion of these populations [122]. The metal tolerant plant species used for bioremediation have developed several mechanisms that allow them to thrive in these contaminated environments and to accumulate high concentrations of specific metals in the aboveground tissue. Among these, both enzymatic and non-enzymatic molecular mechanisms have been described. Heavy-metal stressed plants may protect themselves from reactive oxygen species thought the production of antioxidant enzymes or scavenger compounds [123]. Recently, PGP bacteria, including rhizobia, have been shown to reduce the toxicity of plant exposure to heavy metals [124,125]. Heavy-metal contaminated soil remediation can be performed with different strategies [126]. Plants able to decrease the mobility and/or the bioavailability of metals can be used in both phytostabilization and phytoimmobilization to prevent their leaching into ground water or their entry into the food chain. Mechanisms involved in this process include adsorption by roots, and the precipitation and complexation of the metals in the root zone [127]. Phytovolatilization involves the conversion of a metal (i.e., Hg as the mercuric ion) into the volatile form and its release into the atmosphere through the stomata [128]. However, the most important phytoremediation approach for removal of metals and metalloids from contaminated soils, water, and sediments is phytoextraction [129–131]. The main contribution of rhizobia towards phytostabilization and phytoimmobilization is plant growth enhancement [132]. Bacteria that nodulate their hosts may increase metal accumulation in root nodules, while those that remain in the rhizosphere would reduce metal toxicity locally by precipitation, chelation, immobilization, and biosorption. The nodule itself has an important role in metal-resistance: Once the symbiosis is established, nodules could serve as storage areas that provide plants an extra place to stock metals and reduce the risk of direct exposure [133].

Accelerating the phytoremediation of metalliferous soils by increasing mobilization and phytoextraction of heavy metals though the metabolic activity of rhizobia is a well-known practice [37,131,134–137]. Currently, the most studied metal resistance mechanisms in microorganisms include metal exclusion, protein binding-mediated extra- and intra-cellular sequestration, enzymatic detoxification, active transport of the metal, passive tolerance, and reduction in metal sensitivity of the cellular targets [138,139]. Bacteria can also contribute to phytoremediation through the production of extracellular polymeric substances. For example, studies of the interaction between metals and extracellular polymeric substances demonstrated that biosorption can reduce heavy metal contamination of wastewater systems [140,141]. Unlike salt-tolerant bacteria, there have been numerous studies on the use of bacteria isolated from metal contaminated soil as inoculants to promote plant growth in contaminated

environments [37,48]. Although not all rhizobia are intrinsically tolerant to metals, metal-tolerant strains of taxonomically diverse rhizobia have been isolated from various plants in heavy metal contaminated environments [37,142]. Metal resistance determinants provide protection for rhizobia to survive and maintain effective nodulation of legumes, allowing them to play a role in promoting plant growth. In addition, the existence of a symbiotic relationship may provide protection for the survival of rhizobia in soils with elevated metal concentrations [143,144].

Arsenic toxicity, and the oxidative damage that it produces in cells through the overproduction of reactive oxygen species, affects DNA, proteins, and lipids. This provokes chlorosis, necrosis, delays in flowering, and a reduction in yield [145]. PGP rhizobia may play a beneficial role in protecting plants from arsenic contamination. This can be accomplished by stimulating the antioxidant enzymatic activities in plants, and stabilizing heavy metals and metalloids thereby reducing their accumulation in aerial organs [146–148]. For this reason, the use of PGP rhizobia in heavy metal and metalloid contaminated soils should not only promote the growth of the plant but should also immobilize and decrease the concentration of these elements in plant organs to reduce human exposure to toxic concentrations [149,150]

The presence of heavy metals can also influence the results of inoculant treatment of crops. For example, inoculation of soybean plants with two different strains of *Bradyrhizobium*, *B. diazoefficiens* USDA110 and *Bradyrhizobium* sp. Per 3.61, was studied in the Córdoba province of Argentina where arsenic contamination of groundwater is a consistent environmental problem [151]. The results demonstrated that only *B. diazoefficiens* USDA110 could nodulate soybean at the highest tested As(V) concentrations, while *Bradyrhizobium* sp. Per 3.61 was the better symbiont in the presence of low As(V) concentrations as it limited the translocation of the metal to the legume aerial compartments [151]. Numerous studies have also examined the effect of plant inoculation with pairs of PGP bacteria [152–155]. For instance, the co-inoculation of soybean with *B. japonicum* E109 and *Azospirillum brasilense* Az39influenced plant growth and arsenic phytostabilization in arsenic contaminated conditions [156]. Furthermore, it was observed that the indole acetic acid (IAA) produced by *A. brasilense* Az39 had a protective effect on *B. japonicum* E109 when exposed to arsenic [156]. *B. japonicum* strains can use IAA as a carbon source, which seems to serve as a signal to coordinate bacterial behaviour to enhance protection under adverse conditions [157]. The presence of high levels of lead in soil is toxic for plants, resulting in chlorosis, blackening of roots, and reduced growth [158]. A study using Brassica juncea showed that the inoculation of autochthonous PGP

rhizobial strains can alleviate the harmful effects of lead exposure. *Sinorhizobium* sp. Pb002 was isolated from the rhizosphere of *B. juncea* grown in Pb-contaminated soil [159]. In a microcosm experiment, the presence of strain Pb002 stimulated biomass formation by *B. juncea* and increased plant survival and lead uptake [159]. Nickel and zinc are essential elements for plant growth; however, excessive amounts of these nutrients can be toxic [160,161]. This toxicity can be alleviated, at least in part, through rhizobium inoculation. For example, inoculation of green gram plants with *Bradyrhizobium* sp. (vigna) RM8 or *Rhizobium* sp. RP5 increased both seed yield and grain protein in the presence of excessive nickel or zinc [162,163]. The presence of Cd in soil can impair plant growth due to a reduction in chlorophyll content and photosynthesis [164]. Moreover, Cd alters the cell redox potential and increases the amount of reactive oxygen species in the cell, which in turn negatively impacts cell membranes and biomolecules [164,165]. *Bradyrhizobium* sp. Yl-6, isolated from *G. max* nodules grown in Chinese Cd-contaminated soil, displayed an ability to increase mineral nutrient (Fe) uptake while reducing Cd accumulation [166].

Researchers have identified potential metalloid stress-adaptation genes in rhizobia and they have investigated their transcriptional responses. Putative nickel adaptation genes were identified using association mapping with 47 symbiotic *Mesorhizobium* strains isolated from either nickel-enriched serpentine soils or nearby non-serpentine soils [167]. The identified genes included several transporters, an opine dehydrogenase, and an exopolysaccharide export protein, among others [166]. Additionally, investigation of the transcriptional response of *S. meliloti* strain CCNWSX0020 upon exposure to copper or zinc stress allowed the identification of several upregulated genes, including four genes (yedYZ, fixH-like, cusA-like, and cueO) whose mutation impaired either early or late symbiotic processes [168]. As for rhizobia colonizing saline and arid environments, genetically modified strains can have different symbiotic abilities. For example, a *M. amorphae* ΔcopA deletion mutant displayed impaired symbiotic capabilities in copper contaminated soils, whereas overexpression of a flavodoxin gene in S. meliloti led to a more efficient symbiosis under cadmium stress [169].

Together, the studies discussed in this review (and summarized in Table 1) highlight how the rhizobial genotype can strongly influence symbiotic effectiveness and the plant response in harsh environments. Fully elucidating the genetic and molecular bases of these phenotypes would lay a strong foundation to aid the development of improved bio-inoculants, either through genetic engineering or the rational selection of optimal wild isolates.

Table 1. Studies of plant growth improvement mediated by rhizobium-inoculants on harsh soil.

Strain	Isolation conditions		Crop	Effect / action mechanism	Reference
	Site	Metal concentration			
***Bradyrhizobium diazoefficiens* USDA110**	Ref. strain	As	Soybean	Limits the metalloid translocation and accumulation to edible parts of the legume	[143]
***Bradyrhizobium* sp. Per 3.61**	Nodules of soybeans	As			
***A. brasilense* Az39**		As	Soybean	Enhanced growth of the plant and phytostabilization of As when co-inoculated	[148]
***B. japonicum* E109**		As			
***Sinorhizobium* sp. Pb002**	Rhizospere of *Brassica juncea*	Cd	*Brassica juncea*	Increasing plant survival and lead uptake	[151]
***Bradyrhizobium* sp. *(vigna)* RM8**	Nodules of greengram	Ni, Zn	Greengram	Increasing number of nodules in the plant, IAA and siderophores production	[155]
***Rhizobium* sp. RP5**					
***Bradyrhizobium* sp.**	Nodules of *Glycine max*	Cd	*Glycine max*	Increasing mineral nutrient uptake (Fe) and reducing cadmium accumulation in plant	[159]
Rhizobium leguminosarum (LR-30), Mesorhizobium ciceri (CR.30 and CR-39) and Rhizobium phaseoli (MR-2)	*Lens culinaris L., Vigna radiata L., Cicer aretinum L.*		Wheat (*Triticum aestivum*)	Improving the growth, biomass and drought stress index through production of catalase, IAA and exopolysaccharides under PEG-6000 simulated drought	[160]

Strain	Isolation conditions	Crop	Effect / Mechanism of action	Reference
***Azospirillum* Sp245**	Surface-sterilized wheat roots of (Brazil) [161]	Lattuce (*L. sativa* L.)	Promoting aerial biomass, higher ascorbic acid content accompanied by a lower oxidation rate, better overall visual quality due to higher chlorophyll content, hue, Chroma, L and lower browning intensity	[162]
Sinorhizobium medicae WSM419	Nodules of *Medicago murex*	*Medicago truncatula*	Delay in stress-induced leaf senescence and abscission and nutrient acquisition during drought stress	[163]
Sinorhizobium meliloti A2 strains	Commercial strain, Eastern Canada [164]	*Medicago sativa* cv Apica, M*edicago sativa* cv Halo	Increasing of shoot/root ratio and shoot water content, higher concentrations of starch and pinitol in nodules	[165]
Sinorhizobium meliloti Rm1521	Ottawa vicinity [166]			
***Rhizobium etli* CE3 overexpressing trehalose-6-phosphate synthase gene**	CFN42 derivate[167], original isolate from *P. vulgaris* nodule	*P. vulgaris* var. Negro Jamapa	Enhanced drought tolerance due to upregulation of genes involved in stress tolerance, carbon and nitrogen metabolism by trehalose	[105]
Co-inoculation		**Crop**	**Effect / Mechanism of action**	**Reference**
Rhizobia	PGPRs			
***Rhizobium phaseoli* M6; M9. *phaseoli* M6 and M9 and PGPR**	*Pseudomonas syringae*, Mk1; *Pseudomonas fluorescens* Mk20; *Pseudomonas fluorescens* biot. G Mk25	*Vigna radiata* L.	Decrease of damaging effect of salinity stress on mung bean growth	[114]

5. Development of Rhizobial Inoculants

The use of rhizobial bioinoculants began in the USA at the end of 19th century, where soil containing naturally-occurring rhizobia was mixed with seeds. Since then, rhizobium inoculation has become a common practice to improve crop production [178]. Since the first marketed rhizobium biofertilizer "Nitragin", which was developed by Nobbe and Hiltner in 1896, rhizobial bioformulations have improved dramatically. From the second half of the 19th century, liquid inoculants formulation [179] moved initially to freeze-dried inoculant lyophilization [180], and then to gel-based products such as polyacrylamide (PER) [181], alginate (AER), or xanthan (XER) [27,182]. Over the last 30 years, a huge number of formulations have been patented and commercialized; examples include vermiculite-based Gold CoatTM *Rhizobium* inoculant [183], liquid seed applied soybean inoculant Cell-Tech® [184], liquid in-furrow inoculant LIFT, and air-dried clay powder for alfalfa Nitragin® Gold [27,184]. In 1997, the first marketing of a genetically engineered *S. meliloti* strain RMBPC-2 was approved [185].

Eventually, additives and cell protectant-based liquid inoculant formulations were developed that increased cell survival through the use of compounds such as the polymer polyvinyl pyrrolidone (PVP), carboxymethyl cellulose (CMC) [186], gum Arabic [142], sodium alginate [182], and glycerol [187]. The choice of inoculant carriers that can promote the long-lasting maintenance and protection of viable microbial cells is a global issue [24,188]. Peat is currently the most common organic carrier material for bioformulation production [189], especially in North and South America, Europe, and Australia [190], although other materials (such as coal, bagasse, coir, dust, and perlite) are also used [191]. Supported by successful in vitro experiments, Albareda and collaborators suggested the use of broth culture media as a rhizobial carrier for soybean cultivation [191]. It was demonstrated that after 3 months of storage, liquid cultures were able to maintain more than 10^9 cfu/mL of *Sinorhizobium fredii* SMH12 and *Bradyrhizobium japonicum* USDA110 [191]. Therefore, aqueous-, oil-, or polymer-based liquid formulations have increased enormously in recent years (for details see [190,192–194]). Additionally, the inclusion of microbial or plant secondary metabolites, such as flavonoids and phytohormones, has become a common practice in bioformulation preparation to improve the efficiency of the inoculants [25,195–197].

6. Inoculant Combinations and Phenomic Strategies for Improving Crop Yield

Screening for rhizobia with high nitrogen fixation rates is performed in many laboratories; however, the use of effectiveness as the sole criterion for rhizobia selection may not always be the most relevant criterion for field applications [75]. Indeed, in the soil the rhizobia have to overcome many different adverse conditions (pH, desiccation, nutrient deficiencies, salinity/alkalinity, extreme temperatures, toxicities) [75,198] and they have to outcompete other rhizobial strains [199,200]. It is thought that there is generally an alignment between the fitness of rhizobia and the fitness of their host plants [201]. However, rhizobia are not vertically transmitted but are instead soil bacteria that colonize plant roots. Therefore, ineffective or less effective rhizobia can become abundant and outcompete more effective strains. Moreover, a single plant can be infected by multiple strains with different nitrogen fixation efficiency [202,203]. Recent data suggested that legumes cannot discriminate between effective and ineffective strains prior to infection [204]. Instead, legumes limit the loss of resources by sanctioning individual nodules containing ineffective strains [205]. However ineffective strains may escape from plant sanctions by co-infecting nodules together with effective strains [202].

One way to overcome these competition issues is to select or create highly competitive strains [3]. Rhizobial symbiosis genes (*nod*, *nif*, and *fix*) are generally located on chromosomal mobile elements or on symbiosis plasmids [206]. Taking advantage of these features, it is possible to create hybrid strains without the insertion of exogenous DNA. For example, a hybrid strain of *S. meliloti* was recently created by moving the pSymA megaplasmid (accounting for nearly 20% of total genome content) from a donor *S. meliloti* strain to an acceptor strain [207]. Interesting, the resulting cis-hybrid strain seemed to exhibit a cultivar-specific improvement in symbiotic properties, compared to the parental strains, in controlled laboratory conditions [207]. Similarly, the transfer of symbiotic plasmids between different *R. leguminosarum* strains improved various measures of symbiotic efficiency in laboratory settings [208–211]. Therefore, genome-wide replicon-based remodeling of bacterial strains, potentially supported with a metabolic modelling framework [212], could be a powerful tool in precision agriculture by creating highly efficient strains depending of the farm/soil features [213]. This “Natural Genome Assisted Breeding” approach, based on the transfer of replicons among different strains, will also prevent the introduction of non-natural genes into the environment (Figure 2).

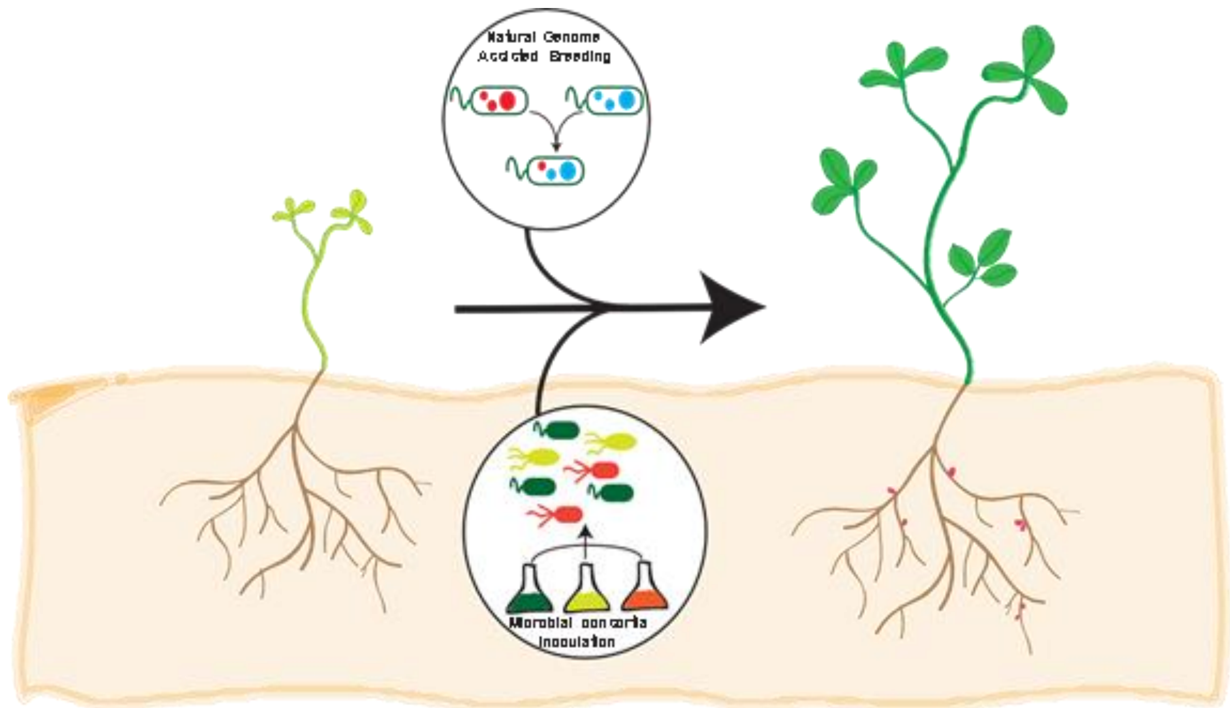


Figure 2. Different approach aimed to promote plant growth: Natural genome assisted breeding as a genomic manipulation strategy, and microbial consortia (multi-strains combination) for the improvement of plant-specific rhizobial inoculants in harsh environments.

One of the major issues related to the application of rhizobium inoculants is the inclusion of other additives (e.g., fungicides, nutrients [193], and fertilizers) that may reduce the viability or effectiveness of rhizobia. A recent trend to overcome this issue is to use microbial consortia instead of single strain biostimulants [214]. Consortia are formed by a combination of bacterial and/or fungal species to cover a broader spectrum of usage and soil conditions [214]. Microbes within a consortium are better able to handle biotic and abiotic stresses as they may work synergistically by exchanging nutrients and removing toxic compounds [215,216]. Using rhizobia in combination with other PGPR may improve their effect; for example, it has been reported that the combination of *Rhizobium* strains with *Bacillus* strains can improve root structure and increase nodule formation in bean, pigeon pea, and soybean (see [217] and references therein). Other well documented examples of mixed inoculants involving rhizobia and PGPR strains that led to improved symbiotic phenotype are: *Rhizobium* with *Bacillus subtilis* and *Bacillus megaterium*; *Rhizobium tropici* with *Chryseobacterium balustinum*, *Bacillus atrophaeus*, and *Burkholderia cepacia*; *Mesorhizobium* with *Pseudomonas*; *Mesorhizobium* in combination with *Azotobacter chroococcum*, *Pseudomonas aeruginosa*, and *Trichoderma harzianum*; and *S. meliloti* with a consortia of *Burkholderia* spp. (see [217] and

references therein). Increased growth promotion could be due to a direct effect on nodulation, the production of phytohormones, or enhanced resistance to crop diseases [216,217]. Additionally, the synergistic interaction between arbuscular mycorrhiza and rhizobia for enhancing crop yield through improving nutrients uptake has also been heavily investigated [217]. Indeed, arbuscular mycorrhiza symbiosis can increase rhizobium nodulation of legumes under control [218] and saline conditions [219]. The consortium can also improve the uptake and transfer of nitrogen in a soybean/maize inter-cropping system [220].

Overall, the use of consortia composed of rhizobia and other PGPR combined with recent advances in rhizobium genomic manipulation could lead to increased inoculum efficiency in field conditions.

7. Concluding Remarks

In recent years, many studies were focused on the development and the optimization of technologies for the improvement of sustainable crop production, in particular in harsh (arid and/or metal-contaminated) environments. The studies were mainly spurred by an increasing requirement for plant proteins, which is due to the increasing worldwide human population and the need to reduce the use of chemical fertilizers. The demand for increasing plant protein production and reducing the use of fertilizers can be accomplished, in part, through the cultivation of legumes; legumes are rich in protein and are able to improve soil fertility through BNF performed by their associated rhizobia. A huge number of studies conducted in the last thirty years have highlighted the ability of rhizobia to colonize particularly harsh soils, and to promote the growth of the leguminous plants to which they are associated. In this review, we summarized the most recent and detailed literature on plant growth promoting rhizobia isolated from (and thus, adapted to) arid and heavy metal contaminated soils, as well as their possible use as inoculants for legume-based agriculture in harsh soils. Despite the current knowledge on the topic, which ranges from genetic to molecular mechanisms, further research should be conducted on the feasibility of plant or soil specific rhizobia-based inoculations. Here, we proposed genetic manipulation strategies, which simulate natural evolution, and strain combination to optimize plant-specific rhizobial inoculants for the improvement of crop yield.

Author Contributions: A.C., conceptualization, original draft preparation, review, editing; A.B., C.F., original draft preparation, literature review sections; G.C.d., original draft preparation, review, editing; F.P., original draft preparation, figures conceptualization; C.V., review, editing. All authors contributed to critically revising the manuscript and gave final approval for publication.

Funding: This research was partially supported by project METATRACK granted by FONDAZIONE CASSA DI RISPARMIO FIRENZE. AC was partially supported by Fondazione Adriano Buzzati-Traverso.

Acknowledgments: The authors are grateful to Alessio Mengoni for critical reading of the manuscript.

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

References

1. Vandenkoornhuyse, P.; Quaiser, A.; Duhamel, M.; Le Van, A.; Dufresne, A. The importance of the microbiome of the plant holobiont. New Phytol. 2015, 206, 1196–1206.

2. Turnbaugh, P.J.; Ley, R.E.; Hamady, M.; Fraser-Liggett, C.M.; Knight, R.; Gordon, J.I. The human microbiome project: Exploring the microbial part of ourselves in a changing world. Nature 2007, 449, 804–810.

3. Checcucci, A.; DiCenzo, G.C.; Bazzicalupo, M.; Mengoni, A. Trade, diplomacy, and warfare: The quest for elite rhizobia inoculant strains. Front. Microbiol. 2017, 8, 2207.

4. Hassani, M.A.; Durán, P.; Hacquard, S. Microbial interactions within the plant holobiont. Microbiome 2018, 6, 58.

5. Sánchez-Cañizares, C.; Jorrín, B.; Poole, P.S.; Tkacz, A. Understanding the holobiont: The interdependence of plants and their microbiome. Curr. Opin. Microbiol. 2017, 38, 188–196.

6. Gans, J.; Wolinsky, M.; Dunbar, J. Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science 2005, 309, 1387–1390.

7. Schloss, P.D.; Handelsman, J. Toward a census of bacteria in soil. PLoS Comput. Biol. 2006, 2, e92.

8. Hartmann, A.; Schmid, M.; van Tuinen, D.; Berg, G. Plant-driven selection of microbes. Plant Soil 2009, 321, 235–257.

9. Pini, F.; East, A.K.; Appia-Ayme, C.; Tomek, J.; Karunakaran, R.; Mendoza-Suárez, M.; Edwards, A.; Terpolilli, J.J.; Roworth, J.; Downie, J.A. Bacterial biosensors for *in vivo* spatiotemporal mapping of root secretion. Plant Physiol. 2017, 174, 1289–1306.

10. Edwards, J.A.; Santos-Medellín, C.M.; Liechty, Z.S.; Nguyen, B.; Lurie, E.; Eason, S.; Phillips, G.; Sundaresan, V. Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. PLoS Biol. 2018, 16, e2003862.

11. Thijs, S.; Sillen, W.; Rineau, F.; Weyens, N.; Vangronsveld, J. Towards an enhanced understanding of plant–microbiome interactions to improve phytoremediation: Engineering the metaorganism. Front. Microbiol. 2016, 7, 341.

12. Berendsen, R.L.; Pieterse, C.M.; Bakker, P.A. The rhizosphere microbiome and plant health. Trends Plant Sci. 2012, 17, 478–486.

13. Mueller, U.G.; Sachs, J.L. Engineering microbiomes to improve plant and animal health. Trends Microbiol. 2015, 23, 606–617.

14. Trivedi, P.; Schenk, P.M.; Wallenstein, M.D.; Singh, B.K. Tiny microbes, big yields: Enhancing food crop production with biological solutions. Microb. Biotechnol. 2017, 10, 999–1003.

15. Pini, F.; Galardini, M.; Bazzicalupo, M.; Mengoni, A. Plant-bacteria association and symbiosis: Are there common genomic traits in Alphaproteobacteria? Genes (Basel) 2011, 2, 1017–1032.

16. Brill, W.J. Genetics of N_2-fixing organisms. In The Biology of Nitrogen Fixation; Quispel, A., Ed.; North-Hol land Pub. Co.: Amsterdam, The Netherlands, 1974.

17. Chi, F.; Shen, S.-H.; Cheng, H.-P.; Jing, Y.-X.; Yanni, Y.G.; Dazzo, F.B. Ascending migration of endophytic rhizobia, from roots to leaves, inside rice plants and assessment of benefits to rice growth physiology. Appl. Environ. Microbiol. 2005, 71, 7271–7278.

18. Vitousek, P.M.; Menge, D.N.L.; Reed, S.C.; Cleveland, C.C. Biological nitrogen fixation: Rates, patterns and ecological controls in terrestrial ecosystems. Philos. Trans. R. Soc. B Biol. Sci. 2013, 368, 20130119.

19. Checcucci, A.; Perrin, E.; Bazzicalupo, M.; Mengoni, A. Genomic diversity and evolution of rhizobia in Microbial Diversity in the Genomic Era; Das, S., Dash, H.R., Eds.; Elsevier Science BV: Amsterdam, The Netherlands, 2019; pp. 37–46.

20. Estrada-de Los Santos, P.; Palmer, M.; Steenkamp, E.T.; Maluk, M.; Beukes, C.; Hirsch, A.M.; James, E.K.; Venter, S.N. *Trinickia dabaoshanensis* sp. nov., a new name for a lost species. Arch. Microbiol. 2019, 1–4.

21. Estrada-de los Santos, P.; Palmer, M.; Chávez-Ramírez, B.; Beukes, C.; Steenkamp, E.; Briscoe, L.; Khan, N.; Maluk, M.; Lafos, M.; Humm, E.; et al. Whole genome analyses suggests that *Burkholderia* sensu lato contains two additional novel genera (*Mycetohabitans* gen. nov., and *Trinickia* gen. nov.): Implications for the Evolution of Diazotrophy and Nodulation in the Burkholderiaceae. Genes 2018, 9, 389.

22. Sprent, J.I.; Ardley, J.; James, E.K. Biogeography of nodulated legumes and their nitrogen-fixing symbionts. New Phytol. 2017, 215, 40–56.

23. Dwivedi, F.; Upadhyaya, K.L.; Galardini, A.; Bazzicalupo, M.; Biondi, M.; Hungria, E.G.; Kaschuk, M.; Blair, G.; Dwivedi, S.L.; Sahrawat, K.L.; et al. Advances in host plant and rhizobium genomics to enhance symbiotic nitrogen fixation in grain legumes. In Advances in Agronomy; Sparks, D.L., Ed.; Elsevier Science BV: Amsterdam, The Netherlands, 2015; Volume 129, pp. 1–116.

24. Galloway, J.N.; Townsend, A.R.; Erisman, J.W.; Bekunda, M.; Cai, Z.; Freney, J.R.; Martinelli, L.A.; Seitzinger, S.P.; Sutton, M.A. Transformation of the nitrogen cycle: Recent trends, questions, and potential solutions. Science 2008, 320, 889–892.

25. Alori, E.T.; Dare, M.O.; Babalola, O.O. Microbial inoculants for soil quality and plant health. In Sustainable Agriculture Reviews; Lichtfouse, E., Ed.; Springer: Cham, Switzerland, 2017; Volume 22, pp. 281–307.

26. Carvajal-Muñoz, J.S.; Carmona-Garcia, C.E. Benefits and limitations of biofertilization in agricultural practices. Livest. Res. Rural Dev. 2012, 24, 1–8.

27. Bhardwaj, D.; Ansari, M.W.; Sahoo, R.K.; Tuteja, N. Biofertilizers function as key player in sustainable agriculture by improving soil fertility, plant tolerance and crop productivity. Microb. Cell Fact. 2014, 13, 66.

28. Arora, N.K.; Verma, M.; Mishra, J. Rhizobial bioformulations: Past, present and future. In Rhizotrophs: Plant Growth Promotion to Bioremediation; Mehnaz, S., Ed.; Springer: Singapore, 2017; Volume 2, pp. 69–99.

29. Tairo, E.V.; Ndakidemi, P.A. Possible benefits of rhizobial inoculation and phosphorus supplementation on nutrition, growth and economic sustainability in grain legumes. Am. J. Res. Commun 2013, 1, 532–556.

30. Smil, V. Enriching the Earth: Fritz Haber, Carl Bosch and the Transformation of World Food Production; The MIT Press: Cambridge, MA, USA, 2001.

31. Herridge, D.F.; Peoples, M.B.; Boddey, R.M. Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil 2008, 311, 1–18.

32. Rajwar, A.; Sahgal, M.; Johri, B.N. Legume–rhizobia symbiosis and interactions in agroecosystems. In Plant Microbe Symbiosis: Fundamentals and Advances; Arora, N., Ed.; Springer: New Delhi, India, 2013; pp. 233–265.

33. Deshwal, V.K.; Singh, S.B.; Kumar, P.; Chubey, A. Rhizobia unique plant growth promoting rhizobacteria: A review. Int. J. Life Sci. 2013, 2, 74–86.

34. Xavier, I.J.; Holloway, G.; Leggett, M. Development of rhizobial inoculant formulations. Crop Manag. 2004, 3.

35. Akibode, C.S.; Maredia, M.K. Global and Regional Trends in Production, Trade and Consumption of Food Legume Crops. 2012. Available online: https://ispc.cgiar.org/sites/default/files/images/Legumetrendsv2.pdf (accessed on 30 April 2019).

36. Pérez-Montaño, F.; Alías-Villegas, C.; Bellogín, R.A.; Del Cerro, P.; Espuny, M.R.; Jiménez-Guerrero, I.; López-Baena, F.J.; Ollero, F.J.; Cubo, T. Plant growth promotion in cereal and leguminous agricultural important plants: From microorganism capacities to crop production. Microbiol. Res. 2014, 169, 325–336.

37. Glick, B.R. Using soil bacteria to facilitate phytoremediation. Biotechnol. Adv. 2010, 28, 367–374.

38. Checcucci, A.; Azzarello, E.; Bazzicalupo, M.; Carlo, A.D.; Emiliani, G.; Mancuso, S.; Spini, G.; Viti, C.; Mengoni, A. Role and regulation of ACC deaminase gene in *Sinorhizobium meliloti*: Is it a symbiotic, rhizospheric or endophytic gene? Front. Genet. 2017, 8, 6.

39. Brígido, C.; Nascimento, F.X.; Duan, J.; Glick, B.R.; Oliveira, S. Expression of an exogenous 1-aminocyclopropane-1-carboxylate deaminase gene in *Mesorhizobium* spp. reduces the negative effects of salt stress in chickpea. FEMS Microbiol. Lett. 2013, 349, 46–53.

40. Djekoun, A.; Planchon, C. Water status effect on dinitrogen fixation and photosynthesis in soybean. Agron. J. 1991, 83, 316–322.

41. Serraj, R.; Sinclair, T.R.; Purcell, L.C. Symbiotic N_2 fixation response to drought. J. Exp. Bot. 1999, 50, 143–155.

42. Elsheik, E.A.E. Effects of salt on rhizobia and bradyrhizobia: A review. Ann. Appl. Biol. 1998, 132, 507–524.

43. Fagorzi, C.; Checcucci, A.; diCenzo, G.; Debiec-Andrzejewska, K.; Dziewit, L.; Pini, F.; Mengoni, A. Harnessing rhizobia to improve heavy-metal phytoremediation by legumes. Genes 2018, 9, 542.

44. Chibuike, G.U.; Obiora, S.C. Heavy metal polluted soils: Effect on plants and bioremediation methods. Appl. Environ. Soil Sci. 2014, 2014, 752708.

45. Lebrazi, S.; Fikri-Benbrahim, K. *Rhizobium*-legume symbioses: Heavy metal effects and principal approaches for bioremediation of contaminated soil. In Legumes for Soil Health and Sustainable Management; Meena, R., Das, A., Yadav, G., Lal, R., Eds.; Springer: Singapore, 2018; pp. 205–233.

46. Mahar, A.; Wang, P.; Ali, A.; Awasthi, M.K.; Lahori, A.H.; Wang, Q.; Li, R.; Zhang, Z. Challenges and opportunities in the phytoremediation of heavy metals contaminated soils: A review. Ecotoxicol. Environ. Saf. 2016, 126, 111–121.

47. Bolan, N.S.; Park, J.H.; Robinson, B.; Naidu, R.; Huh, K.Y. Phytostabilization: A green approach to contaminant containment. Adv. Agron. 2011, 112, 145–204.

48. Kong, Z.; Glick, B.R. The role of plant growth-promoting bacteria in metal phytoremediation. In Advances in Microbial Physiology; Poole, R.K., Ed.; Academic Press: Cambridge, MA, USA, 2017; Volume 71, pp. 97–132.

49. Sturz, A.V.; Nowak, J. Endophytic communities of rhizobacteria and the strategies required to create yield enhancing associations with crops. Appl. Soil Ecol. 2000, 15, 183–190.

50. Theis, K.R.; Dheilly, N.M.; Klassen, J.L.; Brucker, R.M.; John, F.; Bosch, T.C.G.; Cryan, J.F.; Gilbert, S.F.; Goodnight, C.J.; Lloyd, E.A.; et al. Getting the hologenome concept right: An eco—Evolutionary framework for hosts and their microbiomes. Systems 2016, 29, 00028-16.

51. Khan, M.H.; Panda, S.K. Alterations in root lipid peroxidation and antioxidative responses in two rice cultivars under NaCl-salinity stress. Acta Physiol. Plant. 2008, 30, 81.

52. Bouhmouch, I.; Souad-Mouhsine, B.; Brhada, F.; Aurag, J. Influence of host cultivars and *Rhizobium* species on the growth and symbiotic performance of *Phaseolus vulgaris* under salt stress. J. Plant Physiol. 2005, 162, 1103–1113.

53. Zahran, H.H. *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. Microbiol. Mol. Biol. Rev. 1999, 63, 968–989.

54. Cordovilla, M.P.; Ligero, F.; Lluch, C. The effect of salinity on N fixation and assimilation in *Vicia faba*. J. Exp. Bot. 1994, 45, 1483–1488.

55. Abbaspoor, A.; Zabihi, H.R.; Movafegh, S.; Asl, M.H.A. The efficiency of plant growth promoting rhizobacteria (PGPR) on yield and yield components of two varieties of wheat in salinity condition. Am. Eurasian J. Sustain. Agric. 2009, 3, 824–828.

56. Arora, N.K.; Tewari, S.; Singh, S.; Lal, N.; Maheshwari, D.K. PGPR for protection of plant health under saline conditions. In Bacteria in Agrobiology: Stress Management; Maheshwari, D.K., Ed.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 239–258.

57. Vijayan, K.; Chakraborti, S.P.; Ercisli, S.; Ghosh, P.D. NaCl induced morpho-biochemical and anatomical changes in mulberry (*Morus* spp.). Plant Growth Regul. 2008, 56, 61.

58. Drew, M.C.; Hold, P.S.; Picchioni, G.A. Inhibition by NaCl of net CO_2 fixation and yield of cucumber. J. Am. Soc. Hortic. Sci. 1990, 115, 472–477.

59. Cheeseman, J.M. Mechanisms of salinity tolerance in plants. Plant Physiol. 1988, 87, 57.

60. Gupta, B.; Huang, B. Mechanism of salinity tolerance in plants: Physiological, biochemical, and molecular characterization. Int. J. Genom. 2014, 2014, 701596.

61. Cordovilla, M.P.; Ocana, A.; Ligero, F.; Lluch, C. Growth stage response to salinity in symbiosis *Vicia faba-Rhizobium leguminosarum* bv. *viciae*. Plant Physiol. 1995, 14, 105–111.

62. Fagg, C.W.; Stewart, J.L. The value of *Acacia* and *Prosopis* in arid and semi-arid environments. J. Arid Environ. 1994, 27, 3–25.

63. Zhang, X.; Harper, R.; Karsisto, M.; Lindström, K. Diversity of *Rhizobium* bacteria isolated from the root nodules of leguminous trees. Int. J. Syst. Evol. Microbiol. 1991, 41, 104–113.

64. Wahab, A.M.A.; Zahran, H.H. Effects of salt stress on nitrogenase activity and growth of four legumes. Biol. Plant. 1981, 23, 16.

65. Bernstein, L.; Francois, L.E.; Clark, R.A. Interactive effects of salinity and fertility on yields of grains and vegetables 1. Agron. J. 1974, 66, 412–421.

66. Maas, E.V.; Hoffman, G.J. Crop salt tolerance–current assessment. J. Irrig. Drain. Div. 1977, 103, 115–134.

67. Delgado, M.J.; Ligero, F.; Lluch, C. Effects of salt stress on growth and nitrogen fixation by pea, faba-bean, common bean and soybean plants. Soil Biol. Biochem. 1994, 26, 371–376.

68. Sulieman, S.; Tran, L.-S. Asparagine: An amide of particular distinction in the regulation of symbiotic nitrogen fixation of legumes. Crit. Rev. Biotechnol. 2013, 33, 309–327.

69. Ballen, K.G.; Graham, P.H. The role of acid pH in symbiosis between plants and soil organisms. In Handbook of Plant Growth pH as the Master Variable; Rengel, Z., Ed.; Marcel Dekker: New York, NY, USA, 2002; pp. 383–404.

70. Qureshi, M.I.; Muneer, S.; Bashir, H.; Ahmad, J.; Iqbal, M. Nodule physiology and proteomics of stressed legumes. In Advances in Botanical Research; Kader, J.C., Delseny, M., Eds.; Elsevier: Amsterdam, The Netherlands, 2010; Volume 56, pp. 1–48. ISBN 0065-2296.

71. Zahran, H.H. Conditions for successful *Rhizobium*-legume symbiosis in saline environments. Biol. Fertil. Soils 1991, 12, 73–80.

72. Tu, J.C. Effect of salinity on *Rhizobium*-root-hair interaction, nodulation and growth of soybean. Can. J. Plant Sci. 1981, 61, 231–239.

73. Zahran, H.H.; Sprent, J.I. Effects of sodium chloride and polyethylene glycol on root-hair infection and nodulation of *Vicia faba* L. plants by *Rhizobium leguminosarum*. Planta 1986, 167, 303–309.

74. Ikeda, J.; Kobayashi, M.; Takahashi, E. Salt stress increases the respiratory cost of nitrogen fixation. Soil Sci. Plant Nutr. 1992, 38, 51–56.

75. Walsh, K.B. Physiology of the legume nodule and its response to stress. Soil Biol. Biochem. 1995, 27, 637–655.

76. Hungria, M.; Vargas, M.A.T. Environmental factors affecting N_2 fixation in grain legumes in the tropics, with an emphasis on Brazil. Field Crop. Res. 2000, 65, 151–164.

77. Delgado, M.J.; Garrido, J.M.; Ligero, F.; Lluch, C. Nitrogen fixation and carbon metabolism by nodules and bacteroids of pea plants under sodium chloride stress. Physiol. Plant. 1993, 89, 824–829.

78. Zahran, H.H. Rhizobia from wild legumes: Diversity, taxonomy, ecology, nitrogen fixation and biotechnology. J. Biotechnol. 2001, 91, 143–153.

79. Lal, B.; Khanna, S. Selection of salt-tolerant *Rhizobium* isolates of *Acacia nilotica*. World J. Microbiol. Biotechnol. 1994, 10, 637–639.

80. Zahran, H.H.; Räsänen, L.A.; Karsisto, M.; Lindström, K. Alteration of lipopolysaccharide and protein profiles in SDS-PAGE of rhizobia by osmotic and heat stress. World J. Microbiol. Biotechnol. 1994, 10, 100–105.

81. Sakrouhi, I.; Belfquih, M.; Sbabou, L.; Moulin, P.; Bena, G.; Filali-Maltouf, A.; Le Quéré, A. Recovery of symbiotic nitrogen fixing acacia rhizobia from Merzouga desert sand dunes in South East Morocco–Identification of a probable new species of *Ensifer* adapted to stressed environments. Syst. Appl. Microbiol. 2016, 39, 122–131.

82. Gouffi, K.; Pica, N.; Pichereau, V.; Blanco, C. Disaccharides as a new class of nonaccumulated osmoprotectants for *Sinorhizobium meliloti*. Appl. Environ. Microbiol. 1999, 65, 1491–1500.

83. Gouffi, K.; Blanco, C. Is the accumulation of osmoprotectant the unique mechanism involved in bacterial osmoprotection? Int. J. Food Microbiol. 2000, 55, 171–174.

84. Talibart, R.; Jebbar, M.; Gouesbet, G.; Himdi-Kabbab, S.; Wroblewski, H.; Blanco, C.; Bernard, T. Osmoadaptation in rhizobia: Ectoine-induced salt tolerance. J. Bacteriol. 1994, 176, 5210–5217.

85. Arora, N.K.; Singhal, V.; Maheshwari, D.K. Salinity-induced accumulation of poly-β-hydroxybutyrate in rhizobia indicating its role in cell protection. World J. Microbiol. Biotechnol. 2006, 22, 603–606.

86. Breedveld, M.W.; Zevenhuizen, L.P.T.M.; Zehnder, A.J.B. Osmotically-regulated trehalose accumulation and cyclic β-(1, 2)-glucan excretion by *Rhizobium leguminosarum* biovar *trifolii* TA-1. Arch. Microbiol. 1991, 156, 501–506.

87. Ghittoni, N.E.; Bueno, M.A. Changes in the cellular content of trehalose in four peanut rhizobia strains cultured under hypersalinity. Symbiosis 1996, 20, 117–127.

88. Le Rudulier, D.; Bernard, T. Salt tolerance in *Rhizobium*: A possible role for betaines. FEMS Microbiol. Rev. 1986, 2, 67–72.

89. Smith, L.T.; Pocard, J.-A.; Bernard, T.; Le Rudulier, D. Osmotic control of glycine betaine biosynthesis and degradation in *Rhizobium meliloti*. J. Bacteriol. 1988, 170, 3142–3149.

90. Elsheikh, E.A.E.; Wood, M. Salt effects on survival and multiplication of chickpea and soybean rhizobia. Soil Biol. Biochem. 1990, 22, 343–347.

91. Fougere, F.; Le Rudulier, D. Uptake of glycine betaine and its analogues by bacteroids of *Rhizobium meliloti*. Microbiology 1990, 136, 157–163.

92. Hickey, E.W.; Hirshfield, I.N. Low-pH-induced effects on patterns of protein synthesis and on internal pH in *Escherichia coli* and *Salmonella typhimurium*. Appl. Environ. Microbiol. 1990, 56, 1038–1045.

93. Shaharoona, B.; Arshad, M.; Zahir, Z.A.; Khalid, A. Performance of *Pseudomonas* spp. containing ACC-deaminase for improving growth and yield of maize (*Zea mays* L.) in the presence of nitrogenous fertilizer. Soil Biol. Biochem. 2006, 38, 2971–2975.

94. Diagne, N.; Swanson, E.; Pesce, C.; Fall, F.; Diouf, F.; Bakhoum, N.; Fall, D.; Faye, M.N.; Oshone, R.; Simpson, S. Permanent draft genome sequence of *Rhizobium* sp. strain LCM 4573, a salt-tolerant, nitrogen-fixing bacterium isolated from Senegalese soils. Genome Announc. 2017, 5, e00285-17.

95. Ormeño-Orrillo, E.; Gomes, D.F.; del Cerro, P.; Vasconcelos, A.T.R.; Canchaya, C.; Almeida, L.G.P.; Mercante, F.M.; Ollero, F.J.; Megías, M.; Hungria, M. Genome of *Rhizobium leucaenae* strains CFN 299 T and CPAO 29.8: Searching for genes related to a successful symbiotic performance under stressful conditions. BMC Genom. 2016, 17, 534.

96. Molina-Sánchez, M.D.; López-Contreras, J.A.; Toro, N.; Fernández-López, M. Genomic characterization of *Sinorhizobium meliloti* AK21, a wild isolate from the Aral Sea region. Springerplus 2015, 4, 259.

97. Le Quéré, A.; Tak, N.; Gehlot, H.S.; Lavire, C.; Meyer, T.; Chapulliot, D.; Rathi, S.; Sakrouhi, I.; Rocha, G.; Rohmer, M. Genomic characterization of *Ensifer* aridi, a proposed new species of nitrogen-fixing rhizobium recovered from Asian, African and American deserts. BMC Genom. 2017, 18, 85.

98. DiCenzo, G.C.; Zamani, M.; Checcucci, A.; Fondi, M.; Griffitts, J.S.; Finan, T.M.; Mengoni, A. Multidisciplinary approaches for studying *Rhizobium*–legume symbioses. Can. J. Microbiol. 2018, 65, 1–33.

99. Alexandre, A.N.A.; Laranjo, M.; Oliveira, S. Global transcriptional response to heat shock of the legume symbiont *Mesorhizobium loti* MAFF303099 comprises extensive gene downregulation. DNA Res. 2013, 21, 195–206.

100. Laranjo, M.; Alexandre, A.; Oliveira, S. Global transcriptional response to salt shock of the plant microsymbiont *Mesorhizobium loti* MAFF303099. Res. Microbiol. 2017, 168, 55–63.

101. Draghi, W.O.; Del Papa, M.F.; Hellweg, C.; Watt, S.A.; Watt, T.F.; Barsch, A.; Lozano, M.J.; Lagares Jr, A.; Salas, M.E.; López, J.L. A consolidated analysis of the physiologic and

molecular responses induced under acid stress in the legume-symbiont model-soil bacterium *Sinorhizobium meliloti*. Sci. Rep. 2016, 6, 29278.

102. Irar, S.; González, E.M.; Arrese-Igor, C.; Marino, D. A proteomic approach reveals new actors of nodule response to drought in split-root grown pea plants. Physiol. Plant. 2014, 152, 634–645.

103. Larrainzar, E.; Wienkoop, S.; Scherling, C.; Kempa, S.; Ladrera, R.; Arrese-Igor, C.; Weckwerth, W.; González, E.M. Carbon metabolism and bacteroid functioning are involved in the regulation of nitrogen fixation in *Medicago truncatula* under drought and recovery. Mol. Plant-Microbe Interact. 2009, 22, 1565–1576.

104. Larrainzar, E.; Wienkoop, S. A proteomic view on the role of legume symbiotic interactions. Front. Plant Sci. 2017, 8, 1267.

105. Graham, P.H. Stress tolerance in *Rhizobium* and *Bradyrhizobium*, and nodulation under adverse soil conditions. Can. J. Microbiol. 1992, 38, 475–484.

106. Da-Silva, J.R.; Alexandre, A.; Brígido, C.; Oliveira, S. Can stress response genes be used to improve the symbiotic performance of rhizobia. AIMS Microbiol. 2017, 3, 365–382.

107. Suárez, R.; Wong, A.; Ramírez, M.; Barraza, A.; Orozco, M.C.; Cevallos, M.A.; Lara, M.; Hernández, G.; Iturriaga, G. Improvement of drought tolerance and grain yield in common bean by overexpressing trehalose-6-phosphate synthase in rhizobia. Mol. Plant-Microbe Interact. 2008, 21, 958–966.

108. Rodriguez-Valera, F.; Martin-Cuadrado, A.B.; Rodriguez-Brito, B.; Pašic´, L.; Thingstad, T.F.; Rohwer, F.; Mira, A. Explaining microbial population genomics through phage predation. Nat. Rev. Microbiol. 2009, 7, 828–836.

109. Straub, P.F.; Shearer, G.; Reynolds, P.H.S.; Sawyer, S.A.; Kohl, D.H. Effect of disabling bacteroid proline catabolism on the response of soybeans to repeated drought stress. J. Exp. Bot. 1997, 48, 1299–1307.

110. Boscari, A.; Van de Sype, G.; Le Rudulier, D.; Mandon, K. Overexpression of BetS, a *Sinorhizobium meliloti* high-affinity betaine transporter, in bacteroids from *Medicago sativa* nodules sustains nitrogen fixation during early salt stress adaptation. Mol. Plant-Microbe Interact. 2006, 19, 896–903.

111. Nogales, J.; Campos, R.; BenAbdelkhalek, H.; Olivares, J.; Lluch, C.; Sanjuan, J. *Rhizobium tropici* genes involved in free-living salt tolerance are required for the establishment of efficient nitrogen-fixing symbiosis with *Phaseolus vulgaris*. Mol. Plant-Microbe Interact. 2002, 15, 225–232.

112. Vaishnav, A.; Varma, A.; Tuteja, N.; Choudhary, D.K. PGPR-mediated amelioration of crops under salt stress. In Plant-Microbe Interaction: An Approach to Sustainable Agriculture; Choudhary, D.K., Varma, A., Tuteja, N., Eds.; Springer: Singapore, 2016; pp. 205–226.

113. Glick, B.R. Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase. FEMS Microbiol. Lett. 2005, 251, 1–7.

114. Estevez, J.; Dardanelli, M.S.; Megías, M.; Rodríguez-Navarro, D.N. Symbiotic performance of common bean and soybean co-inoculated with rhizobia and *Chryseobacterium balustinum* Aur9 under moderate saline conditions. Symbiosis 2009, 49, 29–36.

115. Aamir, M.; Aslam, A.; Khan, M.Y.; Jamshaid, M.U.; Ahmad, M.; Asghar, H.N.; Zahir, Z.A. Co-inoculation with *Rhizobium* and plant growth promoting rhizobacteria (PGPR) for inducing salinity tolerance in mung bean under field condition of semi-arid climate. Asian J. Agric. Biol. 2013, 1, 17–22.

116. Ahmad, M.; Zahir, Z.A.; Asghar, H.N.; Asghar, M. Inducing salt tolerance in mung bean through coinoculation with rhizobia and plant-growth-promoting rhizobacteria containing 1-aminocyclopropane-1-carboxylate deaminase. Can. J. Microbiol. 2011, 57, 578–589.

117. Etesami, H. Bacterial mediated alleviation of heavy metal stress and decreased accumulation of metals in plant tissues: Mechanisms and future prospects. Ecotoxicol. Environ. Saf. 2018, 147, 175–191.

118. Noll, M.R. Trace elements in terrestrial environments: Biogeochemistry, bioavailability, and risks of metals. J. Environ. Qual. 2003, 32, 374.

119. Ahemad, M.; Khan, M.S. Effect of fungicides on plant growth promoting activities of phosphate solubilizing *Pseudomonas putida* isolated from mustard (*Brassica compestris*) rhizosphere. Chemosphere 2012, 86, 945–950.

120. Pajuelo, E.; Rodríguez-Llorente, I.D.; Dary, M.; Palomares, A.J. Toxic effects of arsenic on *Sinorhizobium-Medicago sativa* symbiotic interaction. Environ. Pollut. 2008, 154, 203–211.

121. Nagajyoti, P.C.; Lee, K.D.; Sreekanth, T.V.M. Heavy metals, occurrence and toxicity for plants: A review. Environ. Chem. Lett. 2010, 8, 199–216.

122. Del Rio, M.; Font, R.; Almela, C.; Velez, D.; Montoro, R.; De Haro Bailon, A. Heavy metals and arsenicuptake by wild vegetation in the Guadiamar river area after the toxic spill of the Aznalcollar mine. J. Biotechnol. 2002, 98, 125–137.

123. Adejumo, S.A. Mechanisms of lead and chromium hyperaccumulation and tolerance in plant. Environtropica 2019, 15, 11–29.

124. Ma, Y.; Oliveira, R.S.; Freitas, H.; Zhang, C. Biochemical and molecular mechanisms of plant-microbe-metal interactions: Relevance for phytoremediation. Front. Plant Sci. 2016, 7, 918.

125. Zubair, M.; Shakir, M.; Ali, Q.; Rani, N.; Fatima, N.; Farooq, S.; Shafiq, S.; Kanwal, N.; Ali, F.; Nasir, I.A. Rhizobacteria and phytoremediation of heavy metals. Environ. Technol. Rev. 2016, 5, 112–119.

126. Sarwar, N.; Imran, M.; Shaheen, M.R.; Ishaque, W.; Kamran, M.A.; Matloob, A.; Rehim, A.; Hussain, S. Phytoremediation strategies for soils contaminated with heavy metals: Modifications and future perspectives. Chemosphere 2017, 171, 710–721.

127. Erakhrumen, A.A. Phytoremediation: An environmentally sound technology for pollution prevention, control and remediation in developing countries. Educ. Res. Rev. 2007, 2, 151–156.

128. Ghosh, M.; Singh, S.P. A review on phytoremediation of heavy metals and utilization of it’s by products. Asian J. Energy Env. 2005, 6, 18.

129. He, Z.L.; Yang, X.E.; Stoffella, P.J. Trace elements in agroecosystems and impacts on the environment. J. Trace Elem. Med. Biol. 2005, 19, 125–140.

130. Seth, C.S. A review on mechanisms of plant tolerance and role of transgenic plants in environmental clean-up. Bot. Rev. 2012, 78, 32–62.

131. Ali, H.; Khan, E.; Sajad, M.A. Phytoremediation of heavy metals—Concepts and applications. Chemosphere 2013, 91, 869–881.

132. Glick, B.R. The enhancement of plant growth by free-living bacteria. Can. J. Microbiol. 1995, 41, 109–117.

133. Hao, X.; Taghavi, S.; Xie, P.; Orbach, M.J.; Alwathnani, H.A.; Rensing, C.; Wei, G. Phytoremediation of Heavy and Transition Metals Aided by Legume-Rhizobia Symbiosis, Int. J. Phytoremediation 2014, 16, 179–202.

134. Rajkumar, M.; Sandhya, S.; Prasad, M.N.; Freitas, H. Perspectives of plant-associated microbes in heavy metal phytoremediation. Biotechnol. Adv. 2012, 30, 1562–1574.

135. Sessitsch, A.; Kuffner, M.; Kidd, P.; Vangronsveld, J.; Wenzel, W.W.; Fallmann, K.; Puschenreiter, M. The role of plant-associated bacteria in the mobilization and phytoextraction of trace elements in contaminated soils. Soil Biol. Biochem. 2013, 60, 182–194.

136. Sharma, R.K.; Archana, G. Cadmium minimization in food crops by cadmium resistant plant growth promoting rhizobacteria. Appl. Soil Ecol. 2016, 107, 66–78.

137. Ullah, A.; Heng, S.; Munis, M.F.H.; Fahad, S.; Yang, X. Phytoremediation of heavy metals assisted by plant growth promoting (PGP) bacteria: A review. Environ. Exp. Bot. 2015, 117, 28–40.

138. Wheaton, G.; Counts, J.; Mukherjee, A.; Kruh, J.; Kelly, R. The confluence of heavy metal biooxidation and heavy metal resistance: Implications for bioleaching by extreme thermoacidophiles. Minerals 2015, 5, 397–451.

139. Bruins, M.R.; Kapil, S.; Oehme, F.W. Microbial resistance to metals in the environment. Ecotoxicol. Environ. Saf. 2000, 45, 198–207.

140. Atieno, M.; Lesueur, D. Opportunities for improved legume inoculants: Enhanced stress tolerance of rhizobia and benefits to agroecosystems. Symbiosis 2018, 77, 191–205.

141. Pal, A.; Paul, A.K. Microbial extracellular polymeric substances: Central elements in heavy metal bioremediation. Indian J. Microbiol. 2008, 48, 49–64.

142. Giller, K.E.; Witter, E.; McGrath, S.P. Heavy metals and soil microbes. Soil Biol. Biochem. 2009, 41, 2031–2037.

143. Martensson, A.; Witter, E. Influence of various soil amendments on nitrogen-fixing soil microorganisms in a long-term field experiment, with special reference to sewage sludge. Soil Biol. Biochem. 1990, 22, 977–982.

144. Obbard, J.P.; Jones, K.C. The effect of heavy metals on dinitrogen fixation by *Rhizobium*-white clover in a range of long-term sewage sludge amended and metal-contaminated soils. Environ. Pollut. 1993, 79, 105–112.

145. Gulz, P.A.; Gupta, S.-K.; Schulin, R. Arsenic accumulation of common plants from contaminated soils. Plant Soil 2005, 272, 337–347.

146. Fatnassi, I.C.; Chiboub, M.; Saadani, O.; Jebara, M.; Jebara, S.H. Phytostabilization of moderate copper contaminated soils using co-inoculation of *Vicia faba* with plant growth promoting bacteria. J. Basic Microbiol. 2015, 55, 303–311.

147. Delgadillo, J.; Lafuente, A.; Doukkali, B.; Redondo-Gómez, S.; Mateos-Naranjo, E.; Caviedes, M.A.; Pajuelo, E.; Rodríguez-Llorente, I.D. Improving legume nodulation and Cu rhizostabilization using genetically modified rhizobia. Environ. Technol. 2015, 36, 1237–1245.

148. Pajuelo, E.; Pérez-Palacios, P.; Romero-Aguilar, A.; Delgadillo, J.; Doukkali, B.; Rodríguez-Llorente, I.D.; Caviedes, M.A. Improving legume–rhizobium symbiosis for copper phytostabilization through genetic manipulation of both symbionts. In Biological Nitrogen Fixation and Beneficial Plant-Microbe Interaction; González-Andrés, F., James, F., Eds.; Springer: Cham, Switzerland, 2016; pp. 183–193.

149. Drewniak, L.; Sklodowska, A. Arsenic-transforming microbes and their role in biomining processes. Environ. Sci. Pollut. Res. 2013, 20, 7728–7739.

150. Pinter, M.I.F.; Salomon, M.V.; Berli, F.; Gil, R.; Bottini, R.; Piccoli, P. Plant growth promoting rhizobacteria alleviate stress by AsIII in grapevine. Agric. Ecosyst. Environ. 2018, 267, 100–108.

151. Bianucci, E.; Godoy, A.; Furlan, A.; Peralta, J.M.; Hernández, L.E.; Carpena-Ruiz, R.O.; Castro, S. Arsenic toxicity in soybean alleviated by a symbiotic species of *Bradyrhizobium*. Symbiosis 2018, 74, 167–176.

152. Fatnassi, I.C.; Chiboub, M.; Saadani, O.; Jebara, M.; Jebara, S.H. Impact of dual inoculation with *Rhizobium* and PGPR on growth and antioxidant status of *Vicia faba* L. under copper stress. C. R. Biol. 2015, 338, 241–254.

153. Dary, M.; Chamber-Pérez, M.A.; Palomares, A.J.; Pajuelo, E. “In situ” phytostabilisation of heavy metal polluted soils using *Lupinus luteus* inoculated with metal resistant plant-growth promoting rhizobacteria. J. Hazard. Mater. 2010, 177, 323–330.

154. Kamaludeen, S.P.B.; Ramasamy, K. Rhizoremediation of metals: Harnessing microbial communities. Indian J. Microbiol. 2008, 48, 80–88.

155. Singh, N.K.; Rai, U.N.; Tewari, A.; Singh, M. Metal accumulation and growth response in *Vigna radiata* L. inoculated with chromate tolerant rhizobacteria and grown on tannery sludge amended soil. Bull. Environ. Contam. Toxicol. 2010, 84, 118–124.

156. Armendariz, A.L.; Talano, M.A.; Nicotra, M.F.O.; Escudero, L.; Breser, M.L.; Porporatto, C.; Agostini, E. Impact of double inoculation with *Bradyrhizobium japonicum* E109 and *Azospirillum brasilense* Az39 on soybean plants grown under arsenic stress. Plant Physiol. Biochem. 2019, 138, 26–35.

157. Spaepen, S.; Vanderleyden, J.; Remans, R. Indole-3-acetic acid in microbial and microorganism-plant signaling. FEMS Microbiol. Rev. 2007, 31, 425–448.

158. Sharma, P.; Dubey, R.S. Lead toxicity in plants. Braz. J. Plant Physiol. 2005, 17, 35–52.

159. Di Gregorio, S.; Barbafieri, M.; Lampis, S.; Sanangelantoni, A.M.; Tassi, E.; Vallini, G. Combined application of Triton X-100 and *Sinorhizobium* sp. Pb002 inoculum for the improvement of lead phytoextraction by *Brassica juncea* in EDTA amended soil. Chemosphere 2006, 63, 293–299.

160. Zhao, F.J.; Lombi, E.; McGrath, S.P. Assessing the potential for zinc and cadmium phytoremediation with the hyperaccumulator *Thlaspi caerulescens*. Plant Soil 2003, 249, 37–43.

161. Seregin, I.V.; Kozhevnikova, A.D. Physiological role of nickel and its toxic effects on higher plants. Russ. J. Plant Physiol. 2006, 53, 257–277.

162. Wani, P.A.; Khan, M.S.; Zaidi, A. Effect of metal-tolerant plant growth-promoting *Rhizobium* on the performance of pea grown in metal-amended soil. Arch. Environ. Contam. Toxicol. 2008, 55, 33–42.

163. Wani, P.A.; Khan, M.S.; Zaidi, A. Effect of metal tolerant plant growth promoting *Bradyrhizobium* sp. (vigna) on growth, symbiosis, seed yield and metal uptake by greengram plants. Chemosphere 2007, 70, 36–45.

164. Hasanuzzaman, M.; Prasad, M.N.V.; Fujita, M. Cadmium Toxicity and Tolerance in Plants: From Physiology to Remediation; Academic Press: Cambridge, MA, USA, 2018; ISBN 0128148659.

165. Yu, R.; Tang, Y.; Liu, C.; Du, X.; Miao, C.; Shi, G. Comparative transcriptomic analysis reveals the roles of ROS scavenging genes in response to cadmium in two pak choi cultivars. Sci. Rep. 2017, 7, 9217.

166. Guo, J.; Chi, J. Effect of Cd-tolerant plant growth-promoting *Rhizobium* on plant growth and Cd uptake by *Lolium multiflorum* Lam. and *Glycine max* (L.) Merr. in Cd-contaminated soil. Plant Soil 2014, 375, 205–214.

167. Porter, S.S.; Chang, P.L.; Conow, C.A.; Dunham, J.P.; Friesen, M.L. Association mapping reveals novel serpentine adaptation gene clusters in a population of symbiotic *Mesorhizobium.* ISME J. 2017, 11, 248–262.

168. Lu, M.; Jiao, S.; Gao, E.; Song, X.; Li, Z.; Hao, X.; Rensing, C.; Wei, G.; Lu, C.M. Transcriptome response to heavy metals in *Sinorhizobium meliloti* CCNWSX0020 reveals new metal eesistance determinants that also promote bioremediation by *Medicago lupulina* in metal-contaminated soil. Appl. Environ. Microbiol. 2017, 83, 1244–1261.

169. Hao, X.; Xie, P.; Zhu, Y.-G.; Taghavi, S.; Wei, G.; Rensing, C. Copper tolerance mechanisms of *Mesorhizobium amorphae* and its role in aiding phytostabilization by *Robinia pseudoacacia* in copper contaminated soil. Environ. Sci. Technol. 2015, 49, 2328–2340.

170. Hussain, M.B.; Zahir, Z.A.; Asghar, H.N.; Asgher, M. Can catalase and exopolysaccharides producing rhizobia ameliorate drought stress in wheat? Int. J. Agric. Biol. 2014, 16, 3–13.

171. Baldani, V.L.D.; Alvarez, M.A.B.; Baldani, J.I.; Döbereiner, J. Establishment of inoculated *Azospirillum* spp. in the rhizosphere and in roots of field grown wheat and sorghum. Plant Soil 1986, 90, 35–46.

172. Fasciglione, G.; Casanovas, E.M.; Quillehauquy, V.; Yommi, A.K.; Goñi, M.G.; Roura, S.I.; Barassi, C.A. *Azospirillum* inoculation effects on growth, product quality and storage life of lettuce plants grown under salt stress. Sci. Hortic. Amst. 2015, 195, 154–162.

173. Staudinger, C.; Mehmeti-Tershani, V.; Gil-Quintana, E.; Gonzalez, E.M.; Hofhansl, F.; Bachmann, G.; Wienkoop, S. Evidence for a rhizobia-induced drought stress response strategy in *Medicago truncatula*. J. Proteom. 2016, 136, 202–213.

174. Bordeleau, L.M; Antoun, H; Lachance, R.-A. Effets des souches de *Rhizobium meliloti* et des coupes successives de la luzerne (*Medicago sativa*) sur la fixation symbiotique d'azote. Can. J. Plant Sci. 1977, 57, 433–439.

175. Bertrand, A.; Dhont, C.; Bipfubusa, M.; Chalifour, F.-P.; Drouin, P.; Beauchamp, C.J. Improving salt stress responses of the symbiosis in alfalfa using salt-tolerant cultivar and rhizobial strain. Appl. Soil Ecol. 2015, 87, 108–117.

176. Bromfield, E.S.P.; Wheatcroft, R.; Barran, L.R. Medium for direct isolation of *Rhizobium meliloti* from soils. Soil Biol. Biochem. 1994, 26, 423–428.

177. Noel, K.D.; Vandenbosch, K.A.; Kulpaca, B. Mutations in *Rhizobium phaseoli* that lead to arrested development of infection threads. J. Bacteriol. 1986, 168, 1392–1401.

178. Smith, R.S. Legume inoculant formulation and application. Can. J. Microbiol. 1992, 38, 485–492.

179. Nobbe, F.; Hiltner, L. Inoculation of the soil for cultivating leguminous plants. US Pat. 1896, 570, 813.

180. Brockwell, J. Inoculation methods for field experimenters and farmers. In Nitrogen Fixation in Legumes; Vincent, J.M., Ed.; Academic Press: New York, NY, USA, 1982; pp. 211–227.

181. Dommergues, Y.R.; Diem, H.G.; Divies, C. Polyacrylamide-entrapped *Rhizobium* as an inoculant for legumes. Appl. Environ. Microbiol. 1979, 37, 779–781.

182. Jung, G.; Mugnier, J.; Diem, H.G.; Dommergues, Y.R. Polymer-entrapped *Rhizobium* as an inoculant for legumes. Plant Soil 1982, 65, 219–231.

183. Paau, A.S.; Graham, L.L.; Bennett, M. Progress in formulation research for PGPR and biocontrol inoculants. Bull. OILB SROP 1991, 14, 8. Available online: http://agris.fao.org/agris-search/search.do?recordID= FR9202869 (accessed on 30 April 2019).

184. Smith, R.S. Inoculant formulations and applications to meet changing needs. In Nitrogen Fixation: Fundamentals and Applications; Tikhonovich, I.A., Provorov, N.A., Romanov, V.I., Newton, W.E., Eds.; Springer: Dordrecht, The Netherlands, 1995; Volume 27, pp. 653–657.

185. Singleton, P.; Keyser, H.; Sande, E. Development and evaluation of liquid inoculants. In Proceedings of the Inoculants and Nitrogen Fixation of Legumes in Vietnam, Hanoi, Vietnam, 17–18 February 2001; pp. 52–66.

186. Rohr, T. Rheological Study of the Mixture Carboxymethylcellulose/Starch and Its Use as a Vehicle for the Bacterial Inoculation. Ph.D. Thesis, Rural Federal University of Rio de Janeiro, Rio De Janeiro, Brazil, 2007.

187. Manikandan, R.; Saravanakumar, D.; Rajendran, L.; Raguchander, T.; Samiyappan, R. Standardization of liquid formulation of *Pseudomonas fluorescens* Pf1 for its efficacy against *Fusarium* wilt of tomato. Biol. Control. 2010, 54, 83–89.

188. Ferreira, E.M.; Castro, I.V. e Residues of the cork industry as carriers for the production of legumes inoculants. Silva. Lusit. 2005, 13, 159–167.

189. Kaljeet, S.; Keyeo, F.; Amir, H. Temperature on survivability of rhizobial inoculant. Asian J. Plant Sci. 2011, 10, 331–337.

190. Bashan, Y.; de-Bashan, L.E.; Prabhu, S.R.; Hernandez, J.-P. Advances in plant growth-promoting bacterial inoculant technology: Formulations and practical perspectives (1998–2013). Plant Soil 2014, 378, 1–33.

191. Albareda, M.; Rodríguez-Navarro, D.N.; Camacho, M.; Temprano, F.J. Alternatives to peat as a carrier for rhizobia inoculants: Solid and liquid formulations. Soil Biol. Biochem. 2008, 40, 2771–2779.

192. Rivera, D.; Obando, M.; Barbosa, H.; Rojas Tapias, D.; Bonilla Buitrago, R. Evaluation of polymers for the liquid rhizobial formulation and their influence in the *Rhizobium*-Cowpea interaction. Univ. Sci. 2014, 19, 265–275.

193. Brahmaprakash, G.P.; Sahu, P.K. Biofertilizers for sustainability. J. Indian Inst. Sci. 2012, 92, 37–62.

194. Vandergheynst, J.; Scher, H.; Guo, H.-Y.; Schultz, D. Water-in-oil emulsions that improve the storage and delivery of the biolarvacide *Lagenidium giganteum*. BioControl 2007, 52, 207–229.

195. Nápoles, M.C.; Carrió, E.G.; Montero, F.; Ferreira, A.; Rossi, A. Role of *Bradyrhizobium japonicum* induced by genistein on soybean stressed by water deficit. Spanish J. Agric. Res. 2009, 7, 665–671.

196. Tank, N.; Saraf, M. Salinity-resistant plant growth promoting rhizobacteria ameliorates sodium chloride stress on tomato plants. J. Plant Interact. 2010, 5, 51–58.

197. Kudoyarova, G.R.; Arkhipova, T.N.; Melent'ev, A.I. Role of bacterial phytohormones in plant growth regulation and their development. In Bacterial Metabolites in Sustainable Agroecosystem; Maeshwari, D., Ed.; Springer: Cham, Switzerland, 2015; pp. 69–86.

198. Giller, K.E.; Murwira, M.S.; Dhliwayo, D.K.C.; Mafongoya, P.L.; Mpepereki, S. Soyabeans and sustainable agriculture in southern Africa. Int. J. Agric. Sustain. 2011, 9, 50–58.

199. Lupwayi, N.Z.; Clayton, G.W.; Rice, W.A. Rhizobial inoculants for legume crops. J. Crop. Improv. 2006, 15, 289–321.

200. Thilakarathna, M.S.; Raizada, M.N. A meta-analysis of the effectiveness of diverse rhizobia inoculants on soybean traits under field conditions. Soil Biol. Biochem. 2017, 105, 177–196.

201. Friesen, M.L. Widespread fitness alignment in the legume—*Rhizobium* symbiosis. New Phytol. 2012, 194, 1096–1111.

202. Checcucci, A.; Azzarello, E.; Bazzicalupo, M.; Galardini, M.; Lagomarsino, A.; Mancuso, S.; Marti, L.; Marzano, M.C.; Mocali, S.; Squartini, A.; et al. Mixed nodule infection in *Sinorhizobium meliloti–Medicago sativa* symbiosis suggest the presence of cheating behavior. Front. Plant Sci. 2016, 7, 835.

203. West, S.A.; Kiers, E.T.; Simms, E.L.; Denison, R.F. Sanctions and mutualism stability: Why do rhizobia fix nitrogen? Proc. R. Soc. B Biol. Sci. 2002, 269, 685–694.

204. Westhoek, A.; Field, E.; Rehling, F.; Mulley, G.; Webb, I.; Poole, P.S.; Turnbull, L.A. Policing the Legume-*Rhizobium* symbiosis: A critical test of partner choice. Sci. Rep. 2017, 7, 1419.

205. Kiers, E.T.; Denison, R.F. Sanctions, cooperation, and the stability of plant-rhizosphere mutualisms. Annu. Rev. Ecol. Evol. Syst. 2008, 39, 215–236.

206. MacLean, A.M.; Finan, T.M.; Sadowsky, M.J. Genomes of the symbiotic nitrogen-fixing bacteria of legumes. Plant Physiol. 2007, 144, 615–622.

207. Checcucci, A.; DiCenzo, G.C.; Ghini, V.; Bazzicalupo, M.; Beker, A.; Decorosi, F.; Dohlemann, J.; Fagorzi, C.; Finan, T.M.; Fondi, M.; et al. Creation and Characterization of a Genomically Hybrid Strain in the Nitrogen-Fixing Symbiotic Bacterium *Sinorhizobium meliloti*. ACS Synth. Biol. 2018, 7, 296483.

208. Brewin, N.J.; Wood, E.A.; Young, J.P.W. Contribution of the symbiotic plasmid to the competitiveness of *Rhizobium leguminosarum*. Microbiology 1983, 129, 2973–2977.

209. Brewin, N.J.; DeJong, T.M.; Phillips, D.A.; Johnston, A.W.B. Co-transfer of determinants for hydrogenase activity and nodulation ability in *Rhizobium leguminosarum*. Nature 1980, 288, 77–79.

210. DeJong, T.M.; Brewin, N.J.; Phillips, D.A. Effects of plasmid content in *Rhizobium leguminosarum* on pea nodule activity and plant growth. Microbiology 1981, 124, 1–7.

211. DiCenzo, G.C.; Checcucci, A.; Bazzicalupo, M.; Mengoni, A.; Viti, C.; Dziewit, L.; Finan, T.M.; Galardini, M.; Fondi, M. Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. Nat. Commun. 2016, 7, 12219.

212. Checcucci, A.; Bazzicalupo, M.; Mengoni, A. Exploiting Nitrogen-Fixing Rhizobial Symbionts Genetic Resources for Improving Phytoremediation of Contaminated Soils. In Enhancing Cleanup of Environmental Pollutants; Anjum, N., Tuteja, N., Eds.; Spinger: Cham, Switzerland, 2017; Volume 1, pp. 275–288.

213. Bradácˇová, K.; Florea, A.S.; Bar-Tal, A.; Minz, D.; Yermiyahu, U.; Shawahna, R.; Kraut-Cohen, J.; Zolti, A.; Erel, R.; Dietel, K. Microbial consortia versus single-strain inoculants: An advantage in PGPM-assisted tomato production? Agronomy 2019, 9, 105.

214. Bashan, Y. Inoculants of plant growth-promoting bacteria for use in agriculture. Biotechnol. Adv. 1998, 16, 729–770.

215. Nadeem, S.M.; Naveed, M.; Zahir, Z.A.; Asghar, H.N. Plant–microbe interactions for sustainable agriculture: Fundamentals and recent advances. In Plant Microbe Symbiosis: Fundamentals and Advances; Arora, N., Ed.; Springer: New Dehli, India, 2013; pp. 51–103.

216. Sekar, J.; Raj, R.; Prabavathy, V.R. Microbial consortial products for sustainable agriculture: Commercialization and regulatory issues in india. In Agriculturally Importantmicroorganisms; Sarma, B., Keswani, C., Eds.; Springer: Singapore, 2016; pp. 107–132.

217. Abd-Alla, M.H.; El-Enany, A.-W.E.; Nafady, N.A.; Khalaf, D.M.; Morsy, F.M. Synergistic interaction of *Rhizobium leguminosarum* bv. *viciae* and arbuscular mycorrhizal fungi as a plant growth promoting biofertilizers for faba bean (*Vicia faba* L.) in alkaline soil. Microbiol. Res. 2014, 169, 49–58.

218. Tajini, F.; Trabelsi, M.; Drevon, J.-J. Co-inoculation with Glomus intraradices and *Rhizobium tropici* CIAT899 increases P use efficiency for N_2 fixation in the common bean (*Phaseolus vulgaris* L.) under P deficiency in hydroaeroponic culture. Symbiosis 2011, 53, 123.

219. Zhu, R.-F.; Tang, F.; Liu, J.; Liu, F.-Q.; Deng, X.-Y.; Chen, J.-S. Co-inoculation of arbuscular mycorrhizae and nitrogen fixing bacteria enhance alfalfa yield under saline conditions. Pakistan J. Bot. 2016, 48, 763–769.

220. Meng, L.; Zhang, A.; Wang, F.; Han, X.; Wang, D.; Li, S. Arbuscular mycorrhizal fungi and *Rhizobium* facilitate nitrogen uptake and transfer in soybean/maize intercropping system. Front. Plant Sci. 2015, 6, 339.

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Chapter 5

Selection of *Rhizobium* strains for inoculation of Lithuanian breeding lines of *Pisum sativum*

Skaidre Supronienе[1]*, Francesca Decorosi[2,3]*, Francesco Pini[2,3,4], Agnese Bellabarba[2,3], Luca Calamai[2], Luciana Giovannetti[2,3], Filippo Bussotti[2], Zydre Kadziuliene[1], Kristyna Razbadauskiene[1], Monika Toleikiene[1], Carlo Viti[2,3]

[1] Institute of Agriculture, Lithuanian Research Centre for Agriculture and Forestry, Instituto al. 1, LT 58344, Akademija, Lithuania.

[2] Department of Agronomy, Food, Environmental and Forestry (DAGRI), University of Florence, Piazzale delle Cascine 28, I-50144, Florence, Italy.

[3] Genexpress Laboratory, Department of Agronomy, Food, Environmental and Forestry (DAGRI), University of Florence, Via della Lastruccia14, I-50019, Sesto Fiorentino, Italy.

[4] Department of Biology, University of Bari Aldo Moro, via Orabona 4, I-70124 Bari, Italy.

Manuscript published on Symbiosis, on 12 January 2021

Personal contribution: performed the plant tests, reviewed the paper

Selection of *Rhizobium* strains for inoculation of Lithuanian breeding lines of *Pisum sativum*

Skaidre Suproniene[1*], Francesca Decorosi[2,3*], Francesco Pini[2,3,4], Agnese Bellabarba[2,3], Luca Calamai[2], Luciana Giovannetti[2,3], Filippo Bussotti[2], Zydre Kadziuliene[1], Kristyna Razbadauskiene[1], Monika Toleikiene[1], Carlo Viti[2,3]

[1] Institute of Agriculture, Lithuanian Research Centre for Agriculture and Forestry, Instituto al. 1, LT-58344, Akademija, Lithuania.

[2] Department of Agronomy, Food, Environmental and Forestry (DAGRI), University of Florence, Piazzale delle Cascine 28, I-50144, Florence, Italy.

[3] Genexpress Laboratory, Department of Agronomy, Food, Environmental and Forestry (DAGRI), University of Florence, Via della Lastruccia14, I-50019, Sesto Fiorentino, Italy.

[4] Department of Biology, University of Bari Aldo Moro, via Orabona 4, I-70124 Bari, Italy.

Corresponding author. E-mail address: skaidre.suproniene@lammc.lt

*These authors contributed equally to this work

Keywords

Inoculant, *Rhizobium*, *Pisum sativum*, Phenotype Microarray

Abstract

Pea (*Pisum sativum*) is one of the most popular legume crops used in agriculture. Because of the high demand and relatively reasonable price, Lithuania has increased the cultivation of this crop and invested in the research of new effective breeding lines in the last years. Rhizobial inoculants contribute to increasing yield in legumes through N_2 fixation. Therefore, the objective of this work was to identify rhizobial strains able to increase the activity of two pea breeding lines (‘DS 3637–2’ and ‘DS 3795–3’) known for high productivity, resistance to biotic and abiotic stresses, and competitiveness in respect to weeds. Six rhizobial strains isolated from pea plants were identified as members of the *Rhizobium leguminosarum* group and phenotypically characterized in depth by Phenotype Microarray (PM). Phenotypic differences observed were linked to their phylogeny. Then, strains were tested for their ability to stimulate the growth of the breeding lines ‘DS 3637–2’ and ‘DS 3795–3’. Reference strain *Rhizobium anhuiense* Z1 and *Rhizobium leguminosarum* sv. *viciae* 14ZE showed the best symbiotic performances with breeding lines ‘DS 3637–2’ and ‘DS 3795–3’, respectively. Based on the obtained results, *R. leguminosarum* sv. *viciae* strain 14ZE appears to be a new effective inoculant of peas.

1. Introduction

Among legume crops used in agriculture, field pea (*Pisum sativum*) is one of the most popular grain legumes, accounting for 16% of world pulse production (Drew et al. 2012). Lithuania, similarly to the other Baltic States, is a country with a longstanding experience in pea cultivation owing to suitable climate and terrain for producing high yielding pulse legumes. Originally, pea crops occupied only about 0.85% of the total arable land (14.9 thousand ha) (Seibutis and Deveikytė 2006). However, the area of pea crops production progressively increased in the last years, covering from 79.4 (2015) to 106.2 (2018) thousand ha of Lithuanian lands (FAOSTAT 2021), and Lithuania became a significant producer of dry beans and dry peas among European Union countries (FAOSTAT 2021).

Rhizobia are a paraphyletic group of well-known bacteria involved in beneficial symbiosis with leguminous plants in a host-specific way (Gage 2004), contributing with a remarkable proportion to the biological nitrogen fixation (BNF) (Udvardi and Poole 2013). In recent years, the exploitation and the improvement of BNF of field pea is a key element for eco-sustainable agriculture (Holt-Giménez and Altieri 2013; Bellabarba et al. 2019; Pastor-Bueis et al. 2019).

It was estimated that more than half of the biologically fixed nitrogen worldwide is yielded by *rhizobium*-legume symbioses (Smil 2004). Moreover, concerning pea crops, BNF can provide over 80% of the nitrogen content in pea plants, while an average of 25–35 kg/ha of nitrogen is introduced in the soil, depending on the tillage system (Ruisi et al. 2012). Consequently, the employment of rhizobial strains as inoculants able to replace commonly used N-fertilizers has become of primary importance and a goal to be achieved for economic and environmental implications (Baset Mia and Shamsuddin 2010; Bhardwaj et al. 2014). Therefore, the isolation and phenotypic screening of rhizobial strains to select as bioinoculants is a crucial strategy. Such screening is also important because the effective symbiosis can be affected by many other factors, like host-specificity and the ability of selected stain to compete with the local rhizobia (Abi-Ghanem et al. 2013; Bourion et al. 2018). The benefits provided for the same legume species by the different rhizobial strains at a given location can vary up to 10-fold (Denison and Kiers 2004).

The *Rhizobium* genus includes more than 90 species; with in this genus, the rhizobia can nodulate different leguminous host plants. The most studied species is *R. leguminosarum*, which could be further classified in different symbiovars depending on the host plant it nodulates: *R. leguminosarum* sv. *viciae* infects plants of tribe Viciaea, like pea (*Pisum sativum* L.) and vetches (*Vicia* L.), while clovers (i.e., *Trifolium* L.) are nodulated by *R. leguminosarum* sv. *trifolii* and common beans (*Phaseolus vulgaris* L) are nodulated by the symbiovar *phaesolii* (Rogel et al. 2011). A tribe could then be nodulated by many different *Rhizobium* species, i.e., plants belonging to the Viciaea tribe may establish symbiotic relations with *R. pisi*, *R. fabae*, *R. laguerreae*, *R. lentis*, *R. bangladeshense*, *R. binae*, *R. anhuiense*, and *R. indicum* (Rahi et al. 2020). Symbiotic interactions between pea and rhizobia are highly selective, i.e., the Afghan *Pisum sativum* cv. Afghanistan is nodulated by specific *R. leguminosarum* sv. *viciae* strains present in Afghan and Turkish soils (Mutch and Young 2004).

In the rhizobia-legume symbiosis, each nodule originates from a separate infection (Kumar et al. 2015). Flavonoids and betaines, released by the root of host-plant attract rhizobia present in the rhizosphere and activate the transcription of nodulation (*nod*) genes, which are located on a large symbiosis plasmid or symbiosis islands. The *nod* genes are responsible for the biosynthesis of nodulation signals (Nod factors), which initiate many of the developmental changes in the roots (Gage 2004). Host specificity depends on nod factors, and the polymorphisms of *nodC* gene are commonly used as phylogenetic markers for *Rhizobium* symbiovar identification (Rogel et al. 2011; Peix et al. 2015). Most symbiovars are currently

defined based on the range of legumes they can nodulate, and this is mainly based on the *nodC* gene analysis, whose phylogenies make possible the correct differentiation of these categories within species of all rhizobial genera (Peix et al. 2015).

Strains of the same *Rhizobium* species recovered from the same field may be genetically different; *R. leguminosarum* sv. *trifolii* strains isolated from nodules of clover plants growing in each other's vicinity showed considerable genetic and metabolic variability (Wielbo et al. 2010; Abi-Ghanem et al. 2013). Metabolic properties may be an important trait in determining nodulation efficiency of rhizobia. Indeed, the ability to metabolize a broad range of carbon and energy sources (especially amino acids and their derivatives) may increase the chance for nodulation of a given strain (Wielbo et al. 2007, 2010). Phenotype microarray is not a time-consuming procedure to examine the metabolic profiles of bacteria; therefore, it could be used to screen rhizobial strains in terms of metabolic versatility (Biondi et al. 2009; Fagorzi et al. 2020).

In the frame to increase the productivity of two *Pisum sativum* emerging breeding lines in Lithuania because of their high productivity, resistance to biotic and abiotic stresses, and competitiveness in respect to weeds (unpublished data), rhizobial strains able to establish an efficient symbiosis were selected and characterized to develop efficient legume growing systems to be employed as eco-sustainable agronomic practices.

2. Materials and Methods

2.1. Isolation of rhizobial strains

Rhizobial strains were isolated from nodules of the four most popular pea cultivars ('Ieva DS', 'Casablanca', 'Respect', and 'Lump') and five breeding lines ('DS 3637-2', 'DS 3486-12', 'DS 3751-1', 'DS3783-3', and 'DS 3789-1') at the flowering stage. Plants were grown for two months (until reaching the flowering stage) in the greenhouse, where the length of the illumination was set up to 16 h, the temperature was 20 ± 2°C, and plants were watered according to the need. Soil for plant growth was collected a few days before experiment setup at 15–20 cm depths from an uninoculated experimental field at the LAMMC ZI Institute (55° 23' 50" N, 23° 51' 40" E). According to the world reference base for soil resources (WRB), the loamy textured soil was classified as an Endocalcic-Epihypogleyic Cambisol (IUSS Working Group WRB 2015). Soil chemical properties of samples were a pH of 6.7 and had labile

phosphorus (P_2O_5) 219 mg/kg, labile potassium (K_2O) 230 mg/kg, total nitrogen (N) 5.87 mg/kg, and organic carbon (Corg) 1.53% (Agrochemical Research Laboratory, Kaunas Lithuania). Single healthy-looking pink nodules (large than 2 mm) were selected from a single plant per cultivar/breeding line and excised. Nodules were surfaced sterilized with 95% ethanol for 10 s and sodium hypochlorite solution containing 3.0% of active chlorine for 3 min, washed five times in sterile deionized water, crushed, and re-suspended in 50 μL of sterile deionized water. Crushed nodule suspensions were serial diluted (1/10 in saline solution) and plated on yeast-mannitol agar (YMA) [0.5 g/L K_2HPO_4, 0.2 g/L $MgSO_4$, 0.1 g/L NaCl, 1.0 g/L $CaCO_3$, mannitol 10.0 g/L, yeast extract 1.0 g/L, agar 20 g/L (Vincent 1970)] and incubated at 28°C for 48 h. Strains were purified streaking a single colony on new YMA plates at least five times. Pure strain cultures were stored in 20% (v/v) glycerol at −80°C.

2.2. 16S rDNA, *rec*A, *atp*D and *nod*C PCR amplification and sequence data analysis

For each strain, a single colony was picked from an YMA plate and re-suspended in 50 μL deionized sterile water. The bacterial suspension was maintained at 100°C for 10 min for cellular lysis, cooled in ice for 2 min, and subsequently centrifuged at 10000 × g for 5 min. Amplification reactions were performed in a 20 μL volume, containing 2 μL of the supernatant as a template. The 16S rRNA was amplified with bacterial universal primer pair 27F 5′-(AGAGTTTGATCMTGGCTCAG)-3 and 1387R 5′-(GGGCGGWGTGTACAAGGC)-3′ (Lane 1991). The composition of the reaction mixture was the following: 1X reaction Buffer (10 mM Tris pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$), 200 μM of each dNTP, 0.1 mg/mL of bovine serum albumin, 0.1 μM of each primer, and 0.05 U of Taq DNA polymerase (Thermo Scientific, Massachusetts, U.S.A.). PCR conditions consisted of an initial denaturation at 94°C for 5 min, 30 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 90 s, and a final extension at 72°C for 5 min.

The *nodC* gene was amplified with primers NodCF 5'-(AYGTHGTYGAYGACGGTTC)-3' and NodCI 5'-(CGYGACAGCCANTCKCTATTG)-3' (Laguerre et al. 2001). The *atpD* gene was amplified with primers AtpD273f 5'-(SCTGGGSCGYATCMTGAACGY)-3' and AtpD771r 5'-(GCCGACACTTCCGAACCNGCCTG)-3' (Gaunt et al. 2001). The *recA* gene was amplified with primers RecA6F 5'-(CGKCTSGTAGAGGAYAAATCGGTGGA)-3' and RecA555r 5'-(CGRATCTGGTTGATGAAGATCACCAT)-3' (Gaunt et al. 2001). Amplification was performed as described above with the only exception of primer

concentrations (0.4 µM of each primer). PCR conditions consisted of an initial denaturation at 94°C for 5 min, 30 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 60 s, and a final extension at 72°C for 5 min.

The PCR amplicons were purified with the Genejet PCR purification kit (Thermo Scientific) and subsequently sequenced by Applied Biosystems 3730XL DNA Analyzer using the same primer set used in PCR amplification. All the sequences were analyzed and edited with BioEdit software (Hall 1999). Phylogenetic analyses of 16S rRNA, *atpD*, and *recA* genes included sequences of *Rhizobium* type strains. The sequences of the *nodC* gene were used as queries to identify related genes using the NCBI BLASTn tool (Altschul et al. 1990) and the non-redundant nucleotide database without a limit to sequences from type material. Phylogenetic analyses were performed using the maximum-likelihood method, selecting the best model for each phylogenetic analysis (Saitou and Nei 1987; Tamura et al. 2004) in the MEGA software (ver. X) (Kumar et al. 2018; Stecher et al. 2020). All ambiguous positions were removed for each sequence (complete deletion option). The 16S rRNA gene, *nodC*, *atpD*, and *recA* gene amplicon sequence data are available at GenBank database with the accession numbers MT775515- MT775524 for the 16S rRNA sequences, MT786692- MT786701 for the *nodC* sequences, MW216573- MW216582 for the *atpD* sequences, and MW216563- MW216571 for the *recA* sequences.

2.3. RAPD analysis

Random Amplification of Polymorphic DNA (RAPD) was performed for all the bacterial strains using primers P7 5'-(CCAAGCTGCC)-3' (Kumar et al. 2009) and AP5 5'-(TCCCGCTGCG)-3' (Caccamo et al. 1999). Amplification reactions were performed as described above with the only exception that only one primer was used at a concentration of 1 µM. PCR conditions were the following: 95°C for 5 min, 40 cycles at 95°C for 1 min, 35°C for 1 min, and 72°C for 1 min and a final extension at 72°C for 7 min. Amplification products were separated by gel electrophoresis on 2% agarose gel and visualized under UV illumination after staining with ethidium bromide. The sizes of RAPD fragments were estimated by comparison with the marker (GeneRuler 1kb DNA ladder, Thermo Scientific. RAPD fingerprints obtained with both primer AP5 and P7 were recorded in binary form, i.e., 1 = presence of a band and 0 = absence of a band. The RAPD matrices obtained with both primers (P7 and AP7) were joined for subsequent cluster analysis. Cluster analyses were performed using the Jaccard similarity

coefficient (J) (Sneath and Sokal 1962) and the algorithm UPGMA (unweighted pair-group method, with arithmetic mean) using PAST v3.25 software (Hammer et al. 2001).

2.4. Plant tests

Two *Pisum sativum* breeding lines 'DS 3637-2' and 'DS 3795-3' were used as host plants. Seeds were surface-sterilized for 1 min in 70% ethanol, rinsed with sterile deionized water, immersed for 5 min in sodium hypochlorite solution containing 5% active chlorine, and washed five times in sterile deionized water. Surface sterilized seeds were allowed to germinate on the cover of sterile plastic Petri dishes over a layer of sterile wet Whatman paper at room temperature. After 4 days in the dark seedlings were transferred to plastic pots filled with a sterilized mixture of sand and vermiculite in a 1:3ratio and supplied with 100 mL of Nitrogen-free solution [1mM $CaCl_2 \cdot 2H_20$, 0.1 mM KCl, 0.8 mM $MgSO_4$ $7H_2O$, 10 µM Fe EDTA, 35 µM H_3BO_3, 9 µM $MnCl_2 \cdot 4H_2O$, 0.8 µM·$ZnCl_2$, 0.5 µM $Na_2MoO \cdot 2H_2O$, 0.3 µM $CuSO_4 \cdot 5H_2O$, 3.68 mM KH_2PO_4, and 4 mM Na_2HPO_4, pH = 6.5] (Ramachandran et al. 2011).

The *Rhizobium* strains were grown at 30°C in liquid TY (Beringer 1974) to the late exponential phase (OD_{600} = 0.6– 0.8), pelleted, and re-suspended to an OD_{600} of 0.05 in a 0.8% NaCl solution. 7-day-old seedlings were inoculated with 5×10^7 rhizobial cells. Five replicates for each *Rhizobium* strain per breeding line were tested. Uninoculated plants were used as a negative control. Plants were grown in a greenhouse under artificial light (luminous flux 3000 lmn, color temperature 4000 K) with a 16-h photoperiod at 22 ± 2°C. Every 6–8 days, plants were watered with sterilized water. Twenty-eight days after inoculation, chlorophyll fluorescence measurements were performed. Intact flag leaves of pea plants were adapted to darkness for 30 min using light-withholding clips. Fluorescence was measured with a portable non-modulated fluorimeter Handy PEA (Plant Efficiency Analyser, Hansatech Instruments, Kings Lynn, UK). The maximum quantum yield of primary photochemistry in the dark-adapted state was expressed for each strain *per* cultivar/ breeding line according to the formula: φPo = [Fm-Fo]/Fm = Fv/Fm] (Bussotti et al. 2011).

The harvested plants were furthermore evaluated for shoot length, number of nodes, number of nodules, and dry weight. Statistical analyses were performed using Rstudio (R Core Team 2011). All data were evaluated for normality with the Shapiro-Wilk test. One-way ANOVA and Tukey post hoc tests were performed for the analysis of shoot length, wet weight, and dry weight by using *agricolae* package. For φPo, the number of nodules and number of nodes

Nonparametric Kruskal-Wallis and post hoc Dunn tests were performed using the *FSA* and *rcompanion* packages. A principal component analysis was performed by using PAST v3.25 software (Hammer et al. 2001).

2.5. Phenotype Microarray

Bacterial strains were characterized using PM1-2 for carbon sources, PM3 for nitrogen sources, PM9 for osmolarity, and PM10 for pH. The complete list of assayed compounds can be obtained from http://www.biolog.com/pdf/PM1-PM10.pdf.

Bacterial strains were grown on TY agar medium (Beringer 1974) and incubated at 27°C for 48 h; then, colonies were picked up with a sterile cotton swab and suspended in 15 mL of NaCl 0.8% solution. Cell density was adjusted to OD_{600} = 0.1. Media for PM1 and PM2 and PM3 were obtained using UMS medium [20 mM 4-morpholine ethanesulfonic acid, 0.5 mM K_2HPO_4, 8.5 mM NaCl, 10 mM NH_4Cl, 10 mM glucose, 2 mM $MgSO_4{\cdot}7H_2O$, 0.51 mM $CaCl_2{\cdot}2H_2O$, 0.4 μM biotin, 4.2 μM Ca-pantothenate, 3.0 μM thiamine, 1 μM Na_2EDTA, 40 μM $FeSO_4{\cdot}7H_2O$, 0.6 μM $ZnSO_4{\cdot}7H_2O$, 0.08 μM $CuSO_4{\cdot}5H_2O$, 4.2 μM $CoCl_2{\cdot}6H_2O$, 0.9 μM $MnSO_4{\cdot}4H_2O$, 4 μM H_3BO_3, 0.8 μM $NaMoO_4{\cdot}2H_2O$, pH 6.5] (Pini et al. 2017) with some modifications. The UMSC was obtained from UMS, without the carbon source (glucose) and used for PM1 and PM2, and the UMSN was obtained from UMS without the nitrogen source (ammonium chloride) and used for PM3. The inoculation fluids for PM1 and PM2 were obtained diluting the cellular suspension (OD_{600} = 0.1) 10 times in UMSC and added with 1X Dye Mix A (Biolog, Hayward, California, U.S.A.). The inoculation fluid for PM3 was prepared by diluting the cellular suspension (OD_{600} = 0.1) 10 times in UMSN added with 1X Dye Mix A (Biolog). The inoculation fluids for PM9 and PM10 were prepared by diluiting the cellular suspension (OD_{600} = 0.1) 10 times in TY added to 1X Dye Mix A (Biolog). Finally, the inoculation fluids were dispensed into PM plates (100 μL per well). The PM plates were sealed with Breathe-Easy® gas-permeable membrane (Sigma-Aldrich, St. Louis, Missouri, U.S.A.) to avoid fluid drying and incubated statically at 27°C in an OmniLog Reader (Biolog) for 96 hours. Readings were recorded for 96 hours, and data were analyzed using OmniLog-PM software (Biolog), which generated a time course curve for tetrazolium color formation. Each experiment was performed in duplicate. The data from the OmniLog-PM software (release OM_PM_109M) were filtered using the area of the kinetic curves as a parameter and then transferred to Excel spreadsheets (Microsoft Corporation). For each carbon and nitrogen source,

the average area was calculated (n = 2). Carbon and nitrogen sources were considered to be used by the strains when the average area was 50% higher than the average area detected in the negative control (A01 well). The TMEV software (Saeed et al. 2003) was used to produce heatmaps of the carbon and nitrogen sources utilization, principal-component analysis (PCA) was performed using the PAST v3.25 software (Hammer et al. 2001) to establish the relationships between the phenotype profiles of the strains.

3. Results and discussion

3.1. Isolation and identification of rhizobial strains

Nine rhizobial strains (2ZE, 3ZE, 4ZE, 7ZE, 9ZE, 11ZE, 13ZE, 14ZE, and 15ZE) were isolated from the nodules of different pea cultivars/breeding lines grown in laboratory conditions. The strain Z1 was added because it is traditionally used in the Institute of Agriculture, Lithuanian Research Centre for Agriculture and Forestry (LAMMC ZI) for the inoculation of pea plants in field experiments. Moreover, the Z1 strains lack genotypic and phenotypic characterization.

Based on phylogenetic analysis by 16S rRNA genes, all strains (including strain Z1) were attributed to the *R. leguminosarum* group (Supplementary Fig. S1). The *R. leguminosarum* group includes *R. leguminosarum* sv. *viciae* USDA 2370^T, *R. laguerreae* FB206^T, *R. sophorae* CCBAU 03386^T, *R. anhuiense* CCBAU 23252^T, *R. acidisoli* FH13^T, *R. ruizarguesonis* UMP1133^T, and *R. hidalgonense* FH14^T, which show 99.9% 16S rRNA gene sequence similarity (Rahi et al. 2020).

The 16S rRNA gene sequencing did not allow the differentiation of strains; therefore, we proceeded typing the strains by RAPD and sequencing the *atpD* and *recA* genes, which are commonly used to differentiate species within the *R. leguminosarum* group (Gaunt et al. 2001). All strains showed different RAPD profiles, except for strains 13ZE and 7ZE (Supplementary Fig. S2). The clustering of RAPD profiles divided the strains into two main groups. The first group contained the strains 3ZE, 7ZE, 11ZE, 13ZE, and 15ZE, while the second group included the strains 2ZE, 4ZE, 9ZE, and 14ZE. The reference strain Z1 was not associated with any other cluster in the dendrogram. Phylogenetic analysis on the concatenated sequences of *atpD* (nucleotide sequence from position 352 to 747) and *recA* (nucleotide sequence from position 160 to 504) confirmed RAPD analysis. The group was formed by the 2ZE, 4ZE, 9ZE, and 14ZE clusters with *R. leguminosarum* sv. *viciae* USDA 2370 (genospecies E), while the group formed

by 3ZE, 7ZE, 11ZE, 13ZE, and 15ZE was related to *R. leguminosarum* sv. *viciae* 3841 (genospecies B) (Fig. 1). Indeed, within the *R. leguminosarum* species there have been identified several different genospecies, though initially only 5 genospecies were characterized (Kumar et al. 2015), recently extensive genomic analysis has subdivided *R. leguminosarum* into 18 different genospecies, which include *R. laguerreae*, *R. sophorae*, *R. ruizarguesonis* and *R. indicum* (Young et al. 2021). Strain Z1, which was the outgroup in the RAPD pro-file, clustered with the sister clade *R. anhuiense* (99.87% sequence identity) (Fig. 1).

The polymorphisms of *nodC* genes are commonly used as phylogenetic markers of *Rhizobium* symbiovars (Peix et al. 2015). The phylogenetic tree based on *nodC* gene sequences showed that all the strains, including the reference strains Z1, belong to the symbiovar viciae (Fig. 2). Strains 3ZE, 9ZE, 11ZE, 13ZE, and 15ZE showed a similar *nodC* sequence and cluster with the *nodC* sequence of the type strain *Rhizobium leguminosarum* sv. *viciae* USDA2370^T (from 97.99% to 99.37% of sequence identity). The other sequences are scattered in the phylogenetic tree. The *nodC* sequence of 14ZE is identical (100%) to the *R. pisi* DSM 30132 *nodC* sequence. Strains Z1 and 2ZE showed 99.5% of *nodC* sequence identity. The finding that all the strains belong to the symbiovar viciae is per the previous statement that Fabeae legumes, such as peas and fava beans, form symbiotic nodules with bacteria belonging to different *Rhizobium* species, with *R. leguminosarum* sv. *viciae* being the most common symbiont for several *Pisum sativum* varieties (Andrews and Andrews 2017). Other *Rhizobium* species able to nodulate plants of the Fabeae tribe (also referred to as the Vicieae) are *R. laguerreae*, *R. sophorae*, *R. anhuiense*, *R. ruizarguesonis*, *R. pisi*, *R. fabeae*, and *R. multihospitium*, which show high sequence similarity of their *nodC* gene suggesting past events of horizontal gene transfer between these species (Andrews and Andrews 2017; Zhang et al. 2019; Jorrin et al. 2020).

Based on molecular analyses, six strains were selected for further characterization. Strain Z1 was selected because was the only strain clustering with *R. anhuiense*, and it is the strain commonly used as an inoculant. Strains 2ZE, 9ZE, and 14ZE were chosen as representatives of *R. leguminosarum* genospecies E, and strains 13ZE and 15ZE were representative of *R. leguminosarum* genospecies B.

3.2. Phenotype microarray analysis: utilization of carbon and nitrogen sources, sensitivity to pH and osmolotyes

The ability of rhizobia to establish an effective symbiosis with the target plants depends not

only on the plant-bacterium interactions but also on several environmental factors, which affect nodulation and nitrogen fixation (Zhang and Smith 2002). Effective rhizobial inoculant must survive in the soil adapting to different chemical-physical proprieties (i.e., pH and salinity), and it must win the competition with native rhizobia and other rhizosphere microorganisms. Nutrient contents in the rhizosphere are influenced by plant roots, which exude large quantities of sugars, organic acids, and amino acids (Gaworzewska and Carlile 1982; Pini et al. 2017), thus affecting the growth of rhizosphere-colonizing bacteria (Jaeger et al. 1999).

Adaptation to different environmental/stress conditions is of fundamental importance to increase the chances for nodulation. Therefore, phenotype microarrays were used to screen the ability of the *Rhizobium* strains to use different nutrients, tolerance to osmotic stressors, and pH conditions (Fig. 3 and S3).

The ability of *Rhizobium* strains to use 190 carbon sources and 95 nitrogen sources was tested using PM1 PM2, and PM3 (Biolog) (Fig. 3a, b and S3). The core of substrates, meaning substrates used by all the strains, was 51 for the carbon sources (4 amino acids of 27 tested, 44 carbohydrates of 93 tested, and 3 carboxylic acids of 47). The strains were diverse in their ability to use the remaining 40 carbon sources (9 amino acids, 15 carbohydrates, 13 carboxylic acids, 1 dipeptide, and 2 others). The remaining carbon sources (99) were not used by any of the strains. Strain 2ZE showed the lowest ability to use carbon sources [60 substrates (31.6% of the compounds tested)], while strains Z1 and 15ZE showed a higher ability to metabolize, respectively, 75 (39%) and 77 (40.5%) of the carbon sources tested. The strains showed a different capability to use the different classes of compounds (carbohydrates < amino acids < carboxylic acids). They used from 51.6 to 57.0% of the carbohydrates tested, from 22.2 to 44.4% of the amino acids tested, and from 8.5 to 25.5% of the carboxylic acid tested (Table 1).

For most of the carbohydrates used, we did not observe large differences among the strains. All the strains used carbon sources whose transporters are induced in *R. leguminosarum* sv. *viciae* 3841 during legume colonization: inositol transported by IntA; orbitol, mannitol, and dulcitol transported by MtlE; raffinose, melibiose, and lactose likely transported by a CUT1 (carbohydrate uptake transporter 1) family SBP (Ramachandran et al. 2011). All the strains were also able to use rhamnose, which has an important role in the early stages of the interaction of *R. leguminosarum* with clover plants (Oresnik et al. 1998). Furthermore, strains 2ZE, 9ZE, 13ZE, 14ZE, and 15ZE were able to use glycerol. It has been shown that the *glp* operon, harboring genes for the glycerol catabolism, was strongly induced by pea seed exudate and that mutants of *R. leguminosarum* sv. *viciae* VF39 unable to use glycerol showed lower nodulation

capabilities (Ding et al. 2012). Strain Z1 showed a higher capability to metabolize different carbohydrates, and this was the only strain able to metabolize D-ribose, N-acetyl-D-mannosamine, N-acetyl-neuraminic acid, and N-acetyl-D-glucosaminitol. All strains (but strain 2ZE) were metabolically active on pectin. It has been demonstrated the presence of pectinolytic activity in rhizobia, but pectin supports lower levels of growth relative to other polysaccharides (Mateos et al. 1992; Knee et al. 2001); however, the lack of pectinolytic activity did not alter the nodulation process (Fauvart et al. 2009).

Dicarboxylic acids play the main role in bacteroid metabolism (Ronson and Primrose 1979; Duncan 1981). Plant photosynthate (sucrose) produced in the shoot is converted to dicarboxylates like succinate, fumarate, and malate (Day 1991) and is provided by the plant to the bacteroids as an energy source for N_2 fixation (Poole and Allaway 2000; Lodwig and Poole 2003). Moreover, plants release large amounts of organic acids in the rhizosphere, contributing to shaping soil microbiota and soil acidification. Bacteria able to use the diverse organic acids secreted by plants then have an advantage in root colonization (Ramachandran et al. 2011). The PM analysis showed that strains 13ZE, 14ZE, 15ZE, and Z1 were able to use succinic and tartaric acid, which are both secreted by pea plants (Pini et al. 2017). Tartaric acid seems to be relevant for pea colonization, the expression of the MFS transporter of the tartrate gene (RL0996) of *R. leguminosarum* sv. *viciae* 3841 was upregulated in the pea rhizosphere and a mutation of this gene negatively affected pea colonization (Ramachandran et al. 2011). In the pea secretome, fumaric and malic acid are also present (Pini et al. 2017), but only two strains (15ZE and Z1) were able to use them (Fig. 3a), at least in free-living conditions.

Regarding the nitrogen source utilization, the core was represented by 40 compounds (10 amino acids, 2 inorganic compounds, 5 purines, 5 pyrimidines, 2 amino sugars, 11 dipeptides, and 5 others). There were 21 sources used only by some of the strains, and 34 compounds were not used by any strain. Strain Z1 and strain 15ZE showed the lowest and the highest ability to use nitrogen sources, metabolizing 45 (47.4% of the compounds tested) and 55 (57.9%) of the substrates tested respectively. Of the three inorganic compounds tested, all the strains were able to metabolize ammonium and nitrate, while nitrite was not used by any strain. All strains used the 12 dipeptides tested. They were also able to use quite a large number of amino sugars [from 3 (50.0%) to 5 (83.3%)], purine [from 5 (55.6%) to 6 (66.7%)], pyrimidines [4 (57.1%)] and amino acids [from 13 (39.4%) to 18 (54.5%)] (Table 1).

In terms of carbon and nitrogen utilization, strain Z1 formed an out-group with respect to all the other strains (Fig. S3A). The other strains could be subdivided into two groups: one included

strains 13ZE and 15ZE and the other included strains 2ZE, 9ZE, and 14ZE, reflecting their different phylogenetic affiliations (Fig. 1 and S3A).

Rhizobium strains differ in their ability to use amino acids, and strain 15ZE was the strain with the highest ability to use these compounds both as carbon and nitrogen sources (Table 1). Out of six strains tested, only 15ZE, 14ZE, and 9ZE strains could use homoserine as a carbon source (Fig. 3a). Strains 13ZE and 15ZE had a similar pattern of amino acid utilization, and they were the only strains using L-ornithine and L-pyroglutamic acid. However, strain 13ZE was not able to metabolize L-homoserine as either a carbon or a nitrogen source. The amino acid homoserine is a major component of pea root exudate; it is released from the main root of pea seedlings during the formation of lateral roots (van Egeraat 1975). The utilization of homoserine is considered an important trait for rhizosphere colonization (Johnston et al. 1987; Hynes and O'Connell 1990) and the horizontal gene transfer of the homoserine gene cluster appears to contribute to rhizosphere colonization (Vanderlinde et al. 2014). Few differences were observed for amino sugars; in particular, strain Z1 showed a negative response for D-galactosamine and D-mannosamine but could use their acetylated forms. With respect to the other nitrogen compounds (peptides, purines, and pyrimidines), the pattern of utilization was similar among the strains tested except for xanthine (not used by strain Z1) and adenine (not used by strains 13ZE, 15ZE, and Z1).

The osmolyte sensitivity (PM9) characterization grouped the strains similar to the carbon and nitrogen utilization groups; Z1 formed an out-group, while the other strains were split into two groups with 13ZE and 15ZE separating from 2ZE, 4ZE, and 9ZE (Fig. S3B), also reflecting phylogenetic differences (Fig. 1).

All strains had low tolerance towards NaCl, showing only a very slight activity on the lowest concentration tested (1%). Rhizobia show marked variation in salt tolerance (Abdelmoumen et al. 1999) and tolerance to salinity of rhizobia was found to be correlated to the geographical location of the soil they inhabit (Cardoso et al. 2015). The low tolerance of the six *Rhizobium* strains to NaCl may reflect their adaptation to the native soil characterized by low salinity and rarely being subjected to desiccation. The tested osmoprotectors did not help the strains to counteract osmotic stress induced by 6% NaCl. Although trehalose is generally recognized as an osmoprotectant in *R. leguminosarum* (McIntyre et al. 2007), the strains were not active in the presence of trehalose and NaCl 6%. Nevertheless, it could not be excluded that trehalose may have a role in osmoprotection at lower NaCl concentrations. The strains did not show metabolic activity on potassium chloride (3–6%), sodium sulfate (2– 5%), sodium formate (1–

6%), urea (2–7%), sodium lactate (1– 12%), or sodium benzoate (20–200 mM), while they tolerated 50 mM ammonium sulfate and 10 mM sodium nitrite. All the strains were resistant to the highest concentrations of ethylene glycol tested (20%). Strain Z1 showed higher tolerance to sodium phosphate and sodium nitrate relative to the other strains. Soil nitrate is one of the typical environmental stresses faced by the legume nodules and their symbiotic partner, inhibiting nodule formation and nitrogenase activity (Gibson and Harper 1985; Herridge and Brockwell 1988; Daimon 1999). However, Saxena et al. (1996) were able to isolate *Rhizobium* sp. strains that nodulated and maintain optimum levels of N fixation in the presence of 8 mM nitrate in combination with a common lentil cultivar host.

The metabolic activity of the *Rhizobium* strains over a broad range of pH (3.5–10) was evaluated. A relevant proportion of the world's soils are acidic (pH < 5.5), and acidity is a widespread feature of Lithuanian soil (Ferguson et al. 2013). Rhizobial survival in soil and nodulation is highly influenced by soil pH (Graham et al. 1994). All the tested *Rhizobium* strains had high activity in the range of pH 6–8 (Fig. 3c). Strains 2ZE, 9ZE, and 14ZE tolerated pH 5.5 well, while strains 13ZE and 15ZE were more sensitive to acid pH and showed higher resistance to basic pH. Strain Z1 showed the highest range of pH resistance, being active in both acid and basic conditions. The obtained results agree with the reported low tolerance of *Rhizobium* to acid pH. Although certain strains of *R. tropici* and *R. loti* are highly acid-tolerant, in general rhizobia did not thrive at pH < 5 [67]. Similar to what was observed for our strains, strains of *R. leguminosarum* sv. *viciae* isolated in Lithuanian soils cannot tolerate acid pH lower than 5.5 (Lapinskas 2007). The ability of strain Z1 to tolerate pH 5.5 is likely one feature making this strain an effective symbiont in Lithuanian soil. Similarly, strains 2ZE, 9ZE, and 14ZE could be good inoculant candidates in this kind of soil.

The global analyses of PM data using a principal component analysis (PCA) divided the six strains into three groups; PC1 accounted for 51.9% of the variability, and PC2 accounted for 25.1%. Strain Z1 was alone in the two-dimensional space and was distinctly separated from strains 2ZE, 9ZE, 13ZE, 14ZE, and 15ZE on PC1. Along PC2, the separation of the other five strains into two main groups occurred; the first included 15ZE, 13ZE, and the second included 2ZE, 9ZE, and 14ZE (Fig. 3d). Observed phenotypic differences were mostly related to strain phylogenetic relationships. Strain Z1 (*R. anhuiense*) was characterized by a high ability to use many different carbon sources (in particular carboxylic acids, together with strain 15ZE), while it was the weakest in terms of nitrogen compound utilization. Moreover, it showed tolerance to a high range of pH. The *R. leguminosarum* strains of genospecies E (2ZE, 9ZE, and 14ZE) was

shown to be well adapted to acidic pH (5.5), while strains of *R. leguminosarum* genospecies B (13ZE and 15ZE) were well adapted to basic pH. Strain 15ZE was more versatile in terms of carbon and nitrogen utilization.

3.3. Greenhouse experiment-assessment of nodulation efficiency and plant growth

The effectiveness of the six strains on *Pisum sativum* was evaluated by using two pea-breeding lines ('DS 3637-2' and 'DS 3795-3') that are new promising cultivars in Lithuania for their improved plant height, longer time of vegetation, and high yields (unpublished).

The two breeding lines showed different characters. The inoculated plants of the breeding line 'DS 3637-2' had shorter shoots, lower dry weights, and a higher number of nodules than the inoculated plants of the breeding line 'DS 3795-3', maintaining at the same time a similar number of nodes. The higher shoot length of 'DS 3795-3' should make this breeding line more advantaged in competition with weeds in respect to the shorter 'DS 3637-2' (McDonald 2003).

To value the effectiveness of the six strains in both the breeding lines, inoculated plants were compared with uninoculated plants. Plants inoculated with 9ZE, 14ZE, 15ZE, and Z1 strains had a significant increase in the shoot length compared to the uninoculated plants. Strain 2ZE only increased shoot length in the breeding line 'DS 3795-3' (Fig. 4a, Supplementary Table S1). Strains 14ZE and 15ZE were able to increase the number of nodes in the symbiotic interaction with 'DS 3637-2' and 'DS 3795-3', respectively, while strain Z1 increased the number of nodes in both pea-breeding lines (Fig. 4b, Supplementary Table S2). A significant increase of pea dry weight was only detected in breeding lines 'DS 3795- 3' and 'DS 3637-2' inoculated with the strains 14ZE and Z1, respectively (Fig. 4c, Supplementary Table S1).

The maximum quantum yield of primary photochemistry in the dark-adapted state (φPo = Fv/Fm) is commonly employed as an index/proxy for photosynthetic efficiency (Baker and Rosenqvist 2004; Bussotti et al. 2011). The photosynthetic efficiencies of both breeding lines 'DS 3795-3' and 'DS 3637-2' were significantly enhanced by all strains (Fig. 4d, Supplementary Table S1). This finding was not consistent with the data reported by Zhou et al. (Zhou et al. 2006), who, when analyzing chlorophyll fluorescence parameters in soybean inoculated with *Sinorhizobium fredii* in the absence of nitrogen, found that rhizobial inoculation had no effects. However, chlorophyll a fluorescence activity, measured in field-grown chickpea plants having variable nodulation performance (high nodulating = HN; low nodulating = LN; non- nodulating = NN), was greater in nodulation variants (HN, and LN) than in the NN variants

(Dudeja and Chaudhary 2005). Peng and co-workers (Peng et al. 2002) demonstrated that rhizobial strains were also able to increase single-leaf net photosynthetic rates in non-legumes, showing that the increase of photosynthetic rate could be not strictly linked to nitrogen fixation. All strains induced the formation of nodules in both breeding lines; as expected, no nodules were detected in the negative controls (Fig. 4e). However, higher numbers of nodules were produced by all strains in the symbiotic interaction with breeding lines 'DS 3637-2' compared to breeding line 'DS3795-3' (Fig. 4e).

Overall, these results showed that strains Z1 and 14ZE were more effective at promoting the activity of *P. sativum* than the other strains tested, while 13ZE had the lowest effect on plant attributes. In both breeding lines, Z1 increased shoot length, the number of nodes, and φPo. Strain 14ZE positively affected shoot length as well as φPo in both breeding lines and it increased the dry weight and the number of nodes in 'DS 3795-3' and 'DS 3637-2', respectively. This result suggested that strain 14ZE may be an effective inoculant for *Pisum sativum*, similar to Z1, already known to efficiently stimulate the growth of *Pisum sativum* from practical field experience. Moreover, the effectiveness of 14ZE may be higher than Z1 and dependent on the inoculated breeding line.

A clear connection between the ability to stimulate *P. sativum* and *nodC* gene sequence cannot be found; strains 9ZE, 13ZE, and 15ZE, which shared similar *nodC* gene sequences (Fig. 1), were less effective than 14ZE and Z1. However, strain 2ZE, whose *nodC* sequence is phylogenetically closer to strain Z1 (Fig. 1), did not show effectiveness similar to that observed for Z1. Differences observed could then be linked to other rhizobial features. Indeed, several traits could be related to effectiveness: i) in the process of adhesion where rhizobia attach to root hairs using different mechanisms such as glucomannan at acidic pH, while at basic pH, the mechanism of adhesion is less understood, ii) variations in exopolysaccharide composition affect host specificity, iii) different strains may respond differently to nodule cysteine-rich peptides produced within the nodule to induce bacteroid differentiation, iv) differences could be related to nitrogenase activity, and v) rhizobia have complex pan-genome whose composition may affect strain effectiveness in different ways (Ghosh and Maiti 2016; Poole et al. 2018; Mendoza-Suárez et al. 2020). Phenotype microarray experiments showed that strain Z1 was able to use a higher number of carboxylic acids. Carboxylic acids are the fuel of bacteroids within nodules (in particular, L-malate, fumarate, and succinate), and this may explain the better effectiveness of strain Z1 and the lower effectiveness of strains 2ZE and 9ZE. Strain 15ZE showed a similar phenotype to strain Z1 for carboxylic acid utilization; in

particular, this strain shared with strain Z1 the ability to use malic acid, which was not found in the other strains. Regardless, plants infected with strain 15ZE did not achieve a high yield as did strain Z1. However, the lack of utilization of malic acid or other carboxylic acids in free-living conditions is not conclusive regarding strain metabolic capabilities within nodules. Differences observed for other compounds have less impact on effectiveness; they may affect it indirectly decreasing/increasing rhizobia survivability in soil, which was not investigated in this study.

PCA was performed on the obtained data of plant performances for the two breeding lines subjected to the different inoculants. (Fig. 5). PC1 and PC2 together accounted for 93.12% and 85.35% of the variance in data obtained from breeding lines 'DS 3637-2' and 'DS 3795-3', respectively. There was a distinct separation between plants treated with bacteria and the uninoculated plants along with the PCs for both the breeding lines (Fig. 5a and b, respectively). A similar separation was observable without considering the number of nodules in the variables used (Fig. S4). Plants inoculated with the two most effective strains 14ZE and Z1 were separated from uninoculated plants along PC1 while a smaller separation was observed for plants inoculated with the other strains along with this PC. Plants inoculated with strain 13ZE, the less effective strain in stimulating pea, were the closest to uninoculated plants along PC1 for both breeding lines (Fig. 5). The phenotypic analysis results did not match with *in vitro* plant test experiments, indicating that the phylogenetic and phenotypic differences observed are probably unlinked to symbiotic effectiveness. However, plant experiments were performed with single strains in sterile conditions using vermiculite and nitrogen-free media, which are the optimal conditions for nodulation. Thus, we cannot exclude that in natural conditions, with the presence of indigenous rhizobia, our strains may behave differently. Nevertheless, strain phenotypic differences should be taken into consideration in the translation from *in vitro* experiments to field experiments where soil features may play a major role in successful colonization. In soil, an increased metabolic versatility may confer competitive advantages; however, future studies on the competitive capabilities of these strains are required to infer a possible relation between strain phenotypic features and competitiveness.

4. Conclusions

The formulation of effective inoculants needs not only to take into account the symbiotic efficiency of the rhizobial strains used but also their phenotypic features expressed as

adaptations to different soil conditions. Here, we used a multiphasic approach for the evaluation of the performance of strains by combining molecular characterization and phenotype microarray analysis with classic nodulation experiments. Phenotypic analyses were correlated to the different phylogenetic affiliations of strains, but they did not match with nodulation assays. Overall, the obtained results suggested that careful attention should be paid to inoculant selection because inoculant efficiency depends on the breeding line used. *Rhizobium* strains 14ZE (*R. leguminosarum*) and Z1 (*R. anhuiense*) could be employed as inoculants for the productivity amelioration of breeding lines 'DS 3795-3' and 'DS 3637-2', respectively.

Acknowledgements

This study has received funding from the Research Council of Lithuania (LMTLT) through High level R&D project "Enhancement of the multifunctional properties of legumes in feed and food value chains (SmartLegume), grant number – 01.2.2-LMT-K-718-01-0068

References

Abdelmoumen H, Filali-Maltouf A, Neyra M, Belabed A, Missbah El Idrissi M (1999). Effect of high salts concentrations on the growth of rhizobia and responses to added osmotica. J Appl Microbiol 86: 889–898. https://doi.org/10.1046/j.1365-2672.1999.00727.x

Abi-Ghanem R, Smith JL, Vandemark GJ (2013). Diversity of *Rhizobium leguminosarum* from pea fields in Washington state. ISRN Soil Sci 2013:786030–786037. https://doi.org/10.1155/2013/786030

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410. https://doi. org/10.1016/S0022-2836(05)80360-2

Andrews M, Andrews ME (2017) Specificity in legume-rhizobia symbioses. Int J Mol Sci 18:705. https://doi.org/10.3390/ijms18040705

Baker NR, Rosenqvist E (2004) Applications of chlorophyll fluorescence can improve crop production strategies: an examination of future possibilities. J Exp Bot 403:1607–1621.

https://doi.org/10.1093/ jxb/erh196

Baset Mia MA, Shamsuddin ZH (2010) Nitrogen fixation and transportation by rhizobacteria: a scenario of rice and banana. Int J Bot 6: 235–242. https://doi.org/10.3923/ijb.2010.235.242

Bellabarba A, Fagorzi C, diCenzo GC, Pini F, Viti C, Checcucci A (2019) Deciphering the symbiotic plant microbiome: translating the most recent discoveries on rhizobia for the improvement of agricultural practices in metal-contaminated and high saline lands. Agronomy 9: 529

Beringer JE (1974) R factor transfer in *Rhizobium leguminosarum*. J Gen Microbiol 84:188–198

Bhardwaj D, Ansari MW, Sahoo RK, Tuteja N (2014) Biofertilizers function as key player in sustainable agriculture by improving soil fertility, plant tolerance and crop productivity. Microb Cell Factories 13:66. https://doi.org/10.1186/1475-2859-13-66

Biondi EG, Tatti E, Comparini D, Giuntini E, Mocali S, Giovannetti L, Mengoni A, Bazzicalupo M, Viti C (2009) Metabolic capacity of *Sinorhizobium* (*Ensifer*) *meliloti* strains as determined by phenotype MicroArray analysis. Appl environ Microbiol 75:5396–5404. Doi: AEM.00196-09 [pii]https://doi.org/10.1128/AEM.00196-09

Bourion V, Heulin-Gotty K, Aubert V, Tisseyre P, Chabart-Martinello M, Pervent M, Delaitre C, Vile D, Siol M (2018) Co-inoculation of a pea core-collection with diverse rhizobial strains shows competitiveness for nodulation and efficiency of nitrogen fixation are distinct traits in the interaction. Front Plant Sci 8:2249. https://doi.org/10. 3389/fpls.2017.02249

Bussotti F, Desotgiu R, Cascio C, Marzuoli R, Nali C, Lorenzini G, Salvatori E, Manes F, Schaub M, Strasser RJ (2011) Ozone stress in woody plants assessed with chlorophyll a fluorescence. A critical reassessment of existing data Environ Exp Bot doi 73:19–30. https://doi.org/10.1016/j.envexpbot.2010.10.022

Caccamo D, Di Cello F, Fani R, Gugliandolo C, Maugeri TL (1999) Polyphasic approach to the characterisation of marine luminous bacteria. Res Microbiol 150:221–230. https://doi.org/10.1016/S0923- 2508(99)80039-4

Cardoso P, Freitas R, Figueira E (2015) Salt tolerance of rhizobial populations from contrasting environmental conditions: understanding the implications of climate change. Ecotoxicology 24:143–152. https://doi.org/10.1007/s10646-014-1366-8

Daimon H (1999) Nitrate-induced inhibition of root nodule formation and nitrogenase activity in the peanut (*Arachis hypogaea* L.). Plant Prod Sci 2:81–86. https://doi.org/10.1626/pps.2.81

Day D (1991) Carbon metabolism and compartmentation in nitrogen-fixing legume nodules. Plant Physiol Biochem 29:185–201

Denison RF, Kiers ET (2004) Lifestyle alternatives for rhizobia: mutualism, parasitism, and forgoing symbiosis. FEMS Microbiol Lett 237: 187–193. https://doi.org/10.1016/j.femsle.2004.07.013

Ding H, Yip CB, Geddes BA, Oresnik IJ, Hynes MF (2012) Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. Microbiology 158:1369– 1378. https://doi.org/10.1099/mic.0.057281-0

Drew EA, Denton MD, Sadras VO, Ballard RA (2012) Agronomic and environmental drivers of population size and symbiotic performance of *Rhizobium leguminosarum* bv. *viciae* in Mediterranean-type environments. Crop Pasture Sci 63:467–477. https://doi.org/10.1071/ CP12032

Dudeja SS, Chaudhary P (2005) Fast chlorophyll fluorescence transient and nitrogen fixing ability of chickpea nodulation variants. Photosynthetica 43:253–259. https://doi.org/10.1007/s11099-005- 0041-y

Duncan MJ (1981) Properties of Tn5-induced carbohydrate mutants in *Rhizobium meliloti*. J Gen Microbiol 122:61–67. https://doi.org/10. 1099/00221287-122-1-61

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

FAOSTAT (2021) Food and agriculture Organization of the United Nations (FAO). FAOSTAT statistical database. In: http://www. fao.org/faostat/en/#data/QC. https://search.library.wisc.edu/catalog/ 999890171702121. Accessed 08/02/2021

Fauvart M, Verstraeten N, Dombrecht B, Venmans R, Beullens S, Heusdens C, Michiels J (2009) *Rhizobium etli* HrpW is a pectin-degrading enzyme and differs from phytopathogenic homologues in enzymically crucial tryptophan and glycine residues. Microbiology 155:3045–3054. https://doi.org/10.1099/mic.0.027599-0

Ferguson BJ, Lin MH, Gresshoff PM (2013) Regulation of legume nodulation by acidic growth conditions. Plant Signal Behav 8:e23426. https://doi.org/10.4161/psb.23426

Gage DJ (2004) Infection and invasion of roots by symbiotic, nitrogen- fixing rhizobia during nodulation of temperate legumes. Microbiol Mol Biol Rev 68:280–300. https://doi.org/10.1128/MMBR.68.2. 280-300.200468/2/280

Gaunt MW, Turner SL, Rigottier-Gois L, Lloyd-Macgilp SA, Young JPW (2001) Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. Int J Syst Evol Microbiol 51: 2037–2048. https://doi.org/10.1099/00207713-51-6-2037

Gaworzewska ET, Carlile MJ (1982) Positive chemotaxis of *Rhizobium leguminosarum* and other bacteria towards root exudates from legumes and other plants. Microbiology 128:1179–1188. https://doi. org/10.1099/00221287-128-6-1179

Ghosh PK, Maiti TK (2016) Structure of extracellular polysaccharides (EPS) produced by rhizobia and their functions in legume–bacteria symbiosis: a review. Achiev Life Sci 10:136–143. https://doi.org/ 10.1016/j.als.2016.11.003

Gibson AH, Harper JE (1985) Nitrate effect on nodulation of soybean by *Bradyrhizobium japonicum* 1. Crop Sci 25:497–501. https://doi.org/ 10.2135/cropsci1985.0011183x002500030015x

Graham PH, Draeger KJ, Ferrey ML, Conroy MJ, Hammer BE, Martinez E, Aarons SR, Quinto C (1994) Acid pH tolerance in strains of *Rhizobium* and *Bradyrhizobium*, and initial studies on the basis for acid tolerance of *Rhizobium tropici* UMR1899. Can J Microbiol 40: 198–297. https://doi.org/10.1139/m94-033

Hall TA (1999) BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95–98

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software packages for education and data analysis. Palaeontol Electron 4:XIX–XX

Herridge DF, Brockwell J (1988) Contributions of fixed nitrogen and soil nitrate to the nitrogen economy of irrigated soybean. Soil Biol Biochem 20:711–717. https://doi.org/10.1016/0038-0717(88)90156-3

Holt-Giménez E, Altieri MA (2013) Agroecology, food sovereignty, and the new green revolution. Agroecol Sustain Food Syst 37:90–102.

https://doi.org/10.1080/10440046.2012.716388

Hynes MF, O'Connell MP (1990) Host plant effect on competition among strains of *Rhizobium leguminosarum*. Can J Microbiol 36: 864–869. https://doi.org/10.1139/m90-150

Jaeger CH, Lindow SE, Miller W, Clark E, Firestone MK (1999) Mapping of sugar and amino acid availability in soil around roots with bacterial sensors of sucrose and tryptophan. Appl Environ Microbiol 65:2685–2690. https://doi.org/10.1128/aem.65.6.2685-2690.1999

Johnston AWB, Downie JA, Rossen L, Shearman CA, Firmin JL, Borthakur D, Wood EA, Bradley D, Brewin NJ, Vincent JM, Haselkorn R, Jones DG, Minchin FR, Dobereiner J, Sprent JI (1987) Molecular analysis of the *Rhizobium* genes involved in the induction of nitrogen-fixing nodules on legumes. Philos Trans R Soc Lond Ser B Biol Sci 317:193–207

Jorrin B, Palacios JM, Peix Á, Imperial J (2020) *Rhizobium ruizarguesonis* sp. nov., isolated from nodules of *Pisum sativum* L. Syst Appl Microbiol 43:126090. https://doi.org/10.1016/j.syapm. 2020.126090

Knee EM, Gong FC, Gao M, Teplitski M, Jones AR, Foxworthy A, Mort AJ, Bauer WD (2001) Root mucilage from pea and its utilization by rhizosphere bacteria as a sole carbon source. Mol Plant-Microbe Interact 6:775–784. https://doi.org/10.1094/MPMI.2001.14.6.775

Kumar PP, Antony LVM, Kumar M (2009) Phylogenetic diversity of cultivable bacteria associated with filamentous non-hetrocystous marine cyanobacteria. Biology

Kumar N, Lad G, Giuntini E, Kaye ME, Udomwong P, Jannah Shamsani N, Young JPW, Bailly X (2015) Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum*. Open Biol 5:140133. https://doi.org/10.1098/rsob. 140133

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/ molbev/msy096

Laguerre G, Nour SM, Macheret V, Sanjuan J, Drouin P, Amarger N (2001) Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris*

symbionts. Microbiology 147:981–993. https://doi.org/10. 1099/00221287-147-4-981

Lane DJ (1991) 16S/23S rRNA Sequencing. Nucleic acid Tech Bact Syst 160:781–791. https://doi.org/10.1007/s00227-012-2133-0

Lapinskas EB (2007) The effect of acidity on the distribution and symbiotic efficiency of rhizobia in Lithuanian soils. Eurasian Soil Sci 40:419–425. https://doi.org/10.1134/S1064229307040084

Lodwig EM, Poole PS (2003) Metabolism of *Rhizobium* bacteroids. CRC Crit Rev Plant Sci 22:37–78

Mateos PF, Jimenez-Zurdo JI, Chen J, Squartini AS, Haack SK, Martinez-Molina E, Hubbell DH, Dazzo FB (1992) Cell-associated pectinolytic and cellulolytic enzymes in *Rhizobium leguminosarum* biovar *trifolii*. Appl Environ Microbiol 58:1816– 1822. https://doi.org/10.1128/aem.58.6.1816-1822.1992

McDonald GK (2003) Competitiveness against grass weeds in field pea genotypes. Weed Res 43:48–58. https://doi.org/10.1046/j.1365- 3180.2003.00316.x

McIntyre HJ, Davies H, Hore TA, Miller SH, Dufour JP, Ronson CW (2007) Trehalose biosynthesis in *Rhizobium leguminosarum* by. *trifolii* and its role in desiccation tolerance. Appl Environ Microbiol 73:3984–3992. https://doi.org/10.1128/AEM.00412-07

Mendoza-Suárez MA, Geddes BA, Sánchez-Cañizares C, Ramírez-González RH, Kirchhelle C, Jorrin B, Poole PS (2020) Optimizing *Rhizobium*-legume symbioses by simultaneous measurement of rhizobial competitiveness and N_2 fixation in nodules. Proc Natl Acad Sci 117:9822–9831. https://doi.org/10.1073/pnas.1921225117

Mutch LA, Young JPW (2004) Diversity and specificity of *Rhizobium leguminosarum* biovar *viciae* on wild and cultivated legumes. Mol Ecol 13:2435–2444. https://doi.org/10.1111/j.1365-294X.2004. 02259.x

Oresnik IJ, Pacarynuk LA, O'Brien SAP, Yost CK, Hynes MF (1998) Plasmid-encoded catabolic genes in *Rhizobium leguminosarum* bv. *trifolii*: evidence for a plant-inducible rhamnose locus involved in competition for nodulation. Mol Plant-Microbe Interact 12:1175–1185. https://doi.org/10.1094/MPMI.1998.11.12.1175

Pastor-Bueis R, Sánchez-Cañizares C, James EK, González-Andrés F (2019) Formulation of a highly effective inoculant for common bean based on an autochthonous elite strain of *Rhizobium leguminosarum* bv. *phaseoli*, and genomic-based insights into its agronomic performance. Front Microbiol 10:2724. https://doi.org/10.3389/fmicb. 2019.02724

Peix A, Ramírez-Bahena MH, Velázquez E, Bedmar EJ (2015) Bacterial associations with legumes. CRC Crit Rev Plant Sci 34:17–42. https://doi.org/10.1080/07352689.2014.897899

Peng S, Biswas JC, Ladha JK, Gyaneshwar P, Chen Y (2002) Influence of rhizobial inoculation on photosynthesis and grain yield of rice. Agron J 4:925–029. https://doi.org/10.2134/agronj2002.9250

Pini F, East AK, Appia-Ayme C, Tomek J, Karunakaran R, Mendoza-Suarez M, Edwards A, Terpolilli JJ, Roworth J, Downie JA, Poole PS (2017) Bacterial biosensors for *in vivo* spatiotemporal mapping of root secretion. Plant Physiol 174:1289–1306. https://doi.org/10.1104/pp.16.01302

Poole P, Allaway D (2020) Carbon and nitrogen metabolism in *Rhizobium*. Adv Microb Physiol 43:117–63. doi: 10.1016/s0065- 2911(00)43004-3. PMID: 10907556

Poole PS, Ramachandran V, Terpolilli J (2018) Rhizobia: from saprophytes to endosymbionts. Nat Rev Microbiol 16:291–303. https:// doi.org/10.1038/nrmicro.2017.171

R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rahi P, Giram P, Chaudhari D, DiCenzo GC, Kiran S, Khullar A, Chandel M, Gawari S, Mohan A, Chavan S, Mahajan B (2020) *Rhizobium indicum* sp. nov., isolated from root nodules of pea (*Pisum sativum*) cultivated in the Indian trans-Himalayas. Syst Appl Microbiol 43:126127. https://doi.org/10.1016/j.syapm.2020. 126127

Ramachandran VK, East AK, Karunakaran R, Downie JA, Poole PS (2011) Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biol 12:R106. https://doi.org/10.1186/gb-2011-12- 10-r106

Rogel MA, Ormeño-Orrillo E, Martinez Romero E (2011) Symbiovars in rhizobia reflect bacterial adaptation to legumes. Syst Appl Microbiol 34:96–104. https://doi.org/10.1016/j.syapm.2010.11.015

Ronson CW, Primrose SB (1979) Carbohydrate metabolism in *Rhizobium trifolii*: identification and symbiotic properties of mutants. J Gen Microbiol 112:77–88. https://doi.org/10.1099/00221287-112-1-77

Ruisi P, Giambalvo D, Di Miceli G, Frenda AS, Saia S, Amato G (2012) Tillage effects on yield and nitrogen fixation of legumes in mediterranean conditions. Agron J 104:1459–1466. https://doi.org/10.2134/ agronj2012.0070

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J (2003) TM4: a free, open-source system for microarray data management and analysis. Biotechniques 34:374–378. https://doi.org/10.2144/03342mt01

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Saxena AK, Rathi SK, Tilak KVBR (1996) Selection and evaluation of nitrate-tolerant strains of *Rhizobium leguminosarum* biovar *viceae* specific to the lentil. Biol Fertil Soils 22:126–130. https://doi.org/10. 1007/BF00384443

Seibutis V, Deveikyté I (2006) Pea yield and its components in different crop rotations. Zemdirbyste-Agriculture 93:263–270

Smil V (2004) Enriching the earth: Fritz Haber, Carl Bosch, and the transformation of world food production. MIT Press, Cambridge

Sneath PHA, Sokal RR (1962) Numerical taxonomy. Nature. 193:855– 860. https://doi.org/10.1038/193855a0

Stecher G, Tamura K, Kumar S (2020) Molecular evolutionary genetics analysis (MEGA) for macOS. Mol Biol Evol 37:1237–1239. https:// doi.org/10.1093/molbev/msz312

Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A 101:11030–11035. https://doi.org/10.1073/pnas. 0404206101

Udvardi M, Poole PS (2013) Transport and metabolism in legume-rhizobia symbioses. Annu Rev Plant Biol 64:781–805. https://doi. org/10.1146/annurev-arplant-050312-120235

van Egeraat AWSM (1975) The possible role of homoserine in the development of *Rhizobium leguminosarum* in the rhizosphere of pea seedlings. Plant Soil 42:381–386. https://doi.org/10.1007/ BF00010013

Vanderlinde EM, Hynes MF, Yost CK (2014) Homoserine catabolism by *Rhizobium leguminosarum* bv. *viciae* 3841 requires a plasmid-borne gene cluster that also affects competitiveness for nodulation. Environ Microbiol 16:205–217. https://doi.org/10.1111/1462-2920.12196

Vincent JM (1970) A manual for the practical study of root-nodule bacteria, . I.B.P. Handbook. Blackwell Sci Publ, Oxford

Wielbo J, Marek-Kozaczuk M, Kubik-Komar A, Skorupska A (2007) Increased metabolic potential of *Rhizobium* spp. is associated with bacterial competitiveness. Can J Microbiol 53:957–967. https://doi. org/10.1139/W07-053

Wielbo J, Marek-Kozaczuk M, Mazur A, Kubik-Komar A, Skorupska A (2010) Genetic and metabolic divergence within a *Rhizobium leguminosarum* bv. *trifolii* population recovered from clover nodules. Appl Environ Microbiol 76:4593–4600. https://doi.org/10. 1128/AEM.00667-10

Young JPW, Moeskjær S, Afonin A, Rahi P, Maluk M, James EK, Cavassim MIA, Rashid MH-o, Aserse AA, Perry BJ, Wang ET, Velázquez E, Andronov EE, Tampakaki A, Flores Félix JD, Rivas González R, Youseif SH, Lepetit M, Boivin S, Jorrin B, Kenicer GJ, Peix Á, Hynes MF, Ramírez-Bahena MH, Gulati A, Tian C-F (2021) Defining the *Rhizobium leguminosarum* species complex. Genes 12: 111 https://doi.org/10.3390/genes12010111

Zhang F, Smith DL (2002) Interorganismal signaling in suboptimum environments: the legume-rhizobia symbiosis. Adv Agron 76:125– 161. https://doi.org/10.1016/s0065-2113(02)76004-5

Zhang J, Shang Y, Peng S, Chen W, Wang E, de Lajudie P, Li B, Guo C, Liu C (2019) *Rhizobium sophorae*, *Rhizobium laguerreae*, and two novel *Rhizobium* genospecies associated with *Vicia sativa* L. in Northwest China. Plant Soil 442:113–126. https://doi.org/10.1007/ s11104-019-04168-w

Zhou XJ, Liang Y, Chen H, Shen SH, Jing YX (2006) Effects of rhizobia inoculation and nitrogen fertilization on photosynthetic physiology of soybean. Photosynthetica 44:530–535. https://doi.org/10.1007/ s11099-006-0066-x

Figures

Figure 1. Phylogenetic tree based on the concatenated alignment of *atpD* and *recA* gene sequences inferred using the maximum likelihood and general time reversible model. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the general time reversible model, and then selecting the topology with superior log likelihood value. Letters in circles indicate the different *R. leguminosarum* genospecies. The

tree is drawn to scale, with branch lengths measured in the number of substitutions per site. There were 738 positions in the final dataset.

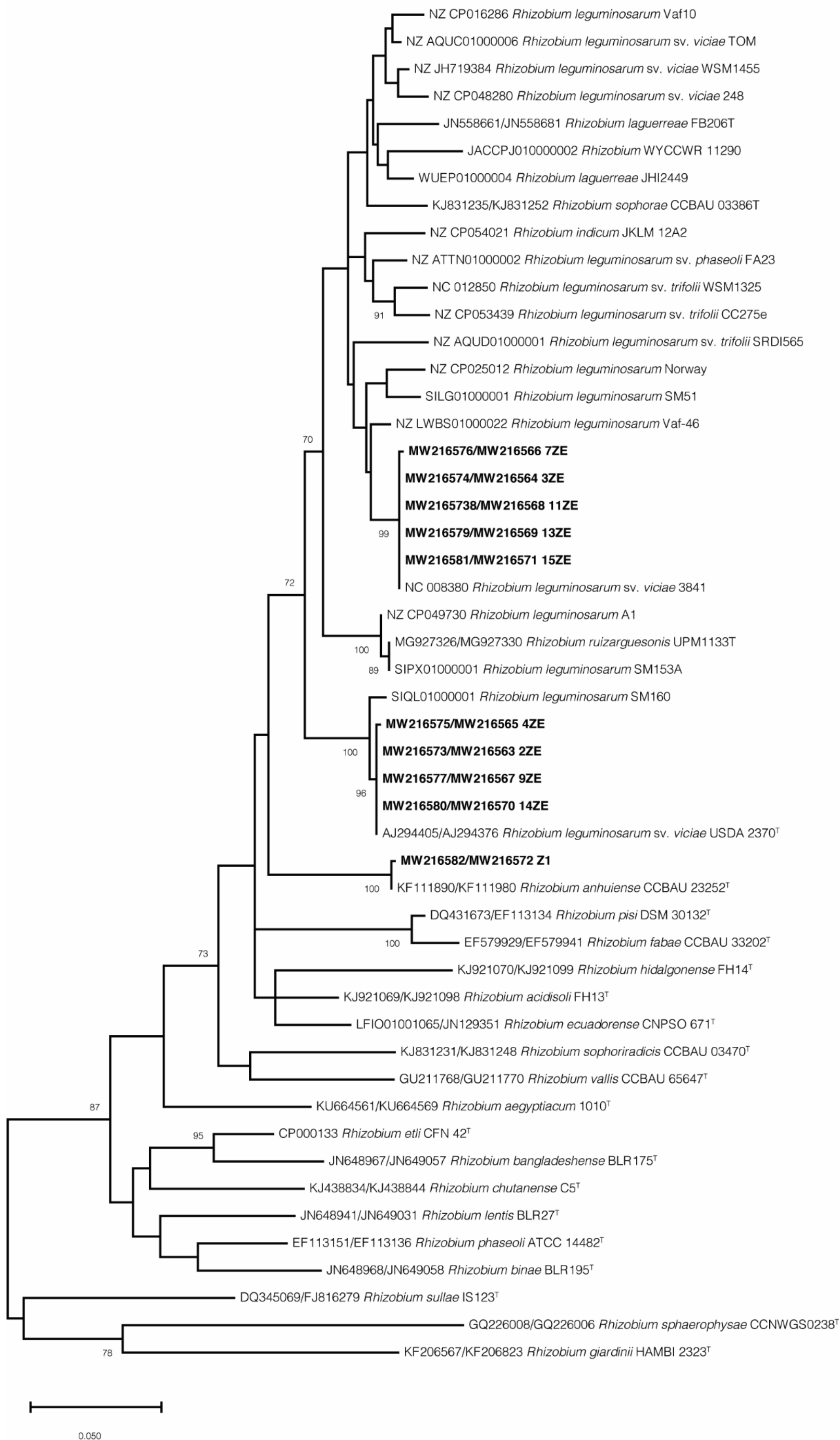
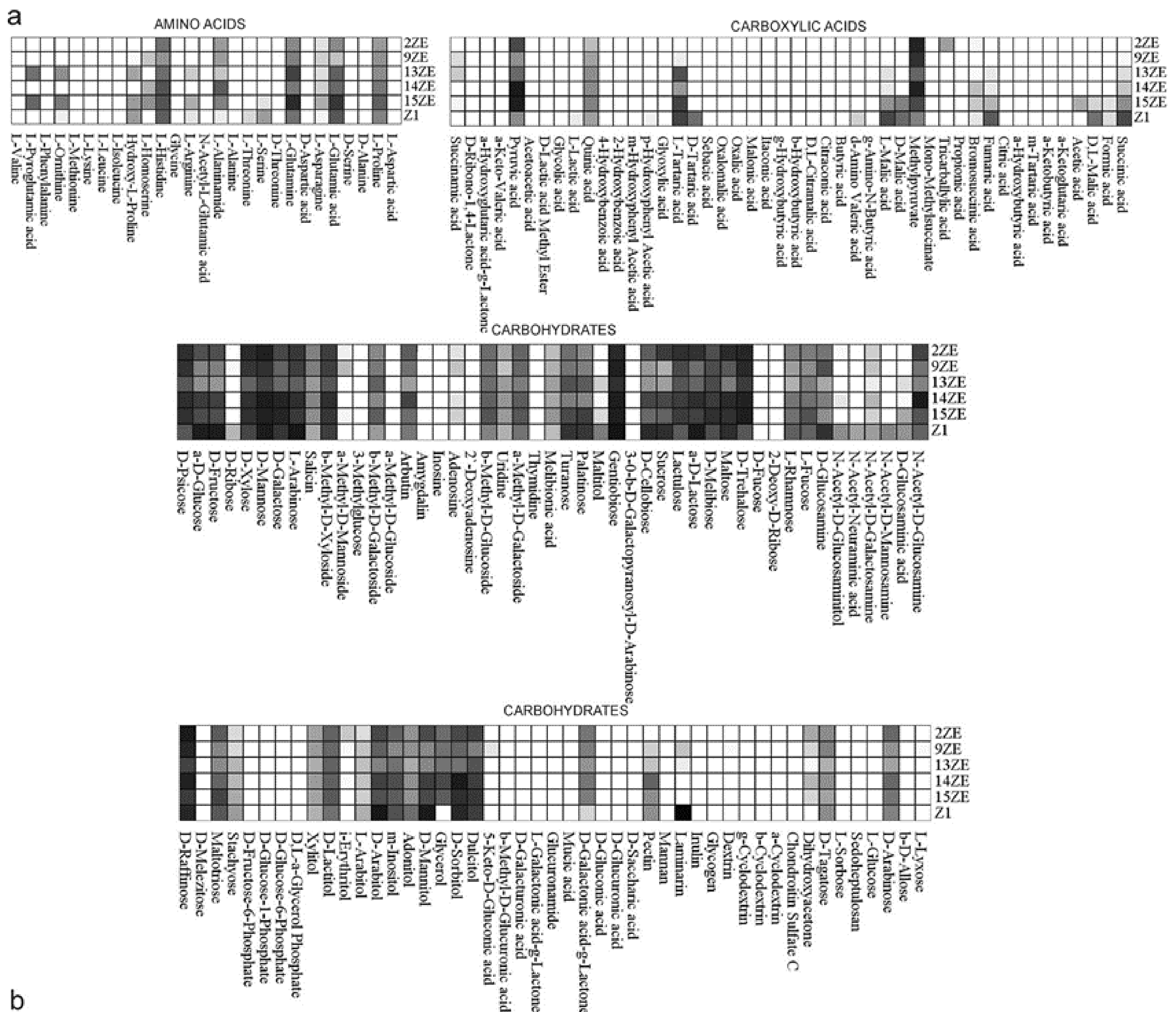

NZ CP016286 Rhizobium leguminosarum Vaf10
NZ AQUC01000006 Rhizobium leguminosarum sv. viciae TOM
NZ JH719384 Rhizobium leguminosarum sv. viciae WSM1455
NZ CP048280 Rhizobium leguminosarum sv. viciae 248
JN558661/JN558681 Rhizobium laguerreae FB206T
JACCPJ010000002 Rhizobium WYCCWR 11290
WUEP01000004 Rhizobium laguerreae JHI2449
KJ831235/KJ831252 Rhizobium sophorae CCBAU 03386T
NZ CP054021 Rhizobium indicum JKLM 12A2
NZ ATTN01000002 Rhizobium leguminosarum sv. phaseoli FA23
NC 012850 Rhizobium leguminosarum sv. trifolii WSM1325
NZ CP053439 Rhizobium leguminosarum sv. trifolii CC275e
NZ AQUD01000001 Rhizobium leguminosarum sv. trifolii SRDI565
NZ CP025012 Rhizobium leguminosarum Norway
SILG01000001 Rhizobium leguminosarum SM51
NZ LWBS01000022 Rhizobium leguminosarum Vaf-46
MW216576/MW216566 7ZE
MW216574/MW216564 3ZE
MW2165738/MW216568 11ZE
MW216579/MW216569 13ZE
MW216581/MW216571 15ZE
NC 008380 Rhizobium leguminosarum sv. viciae 3841
NZ CP049730 Rhizobium leguminosarum A1
MG927326/MG927330 Rhizobium ruizarguesonis UPM1133T
SIPX01000001 Rhizobium leguminosarum SM153A
SIQL01000001 Rhizobium leguminosarum SM160
MW216575/MW216565 4ZE
MW216573/MW216563 2ZE
MW216577/MW216567 9ZE
MW216580/MW216570 14ZE
AJ294405/AJ294376 Rhizobium leguminosarum sv. viciae USDA 2370T
MW216582/MW216572 Z1
KF111890/KF111980 Rhizobium anhuiense CCBAU 23252T
DQ431673/EF113134 Rhizobium pisi DSM 30132T
EF579929/EF579941 Rhizobium fabae CCBAU 33202T
KJ921070/KJ921099 Rhizobium hidalgonense FH14T
KJ921069/KJ921098 Rhizobium acidisoli FH13T
LFIO01001065/JN129351 Rhizobium ecuadorense CNPSO 671T
KJ831231/KJ831248 Rhizobium sophoriradicis CCBAU 03470T
GU211768/GU211770 Rhizobium vallis CCBAU 65647T
KU664561/KU664569 Rhizobium aegyptiacum 1010T
CP000133 Rhizobium etli CFN 42T
JN648967/JN649057 Rhizobium bangladeshense BLR175T
KJ438834/KJ438844 Rhizobium chutanense C5T
JN648941/JN649031 Rhizobium lentis BLR27T
EF113151/EF113136 Rhizobium phaseoli ATCC 14482T
JN648968/JN649058 Rhizobium binae BLR195T
DQ345069/FJ816279 Rhizobium sullae IS123T
GQ226008/GQ226006 Rhizobium sphaerophysae CCNWGS0238T
KF206567/KF206823 Rhizobium giardinii HAMBI 2323T
P
N
J
Q
R
S
O
G
I
K
H
A
M
D
L
B
C
C
C
E
E
91
70
99
72
100
89
100
96
100
100
73
87
95
78
0.050

Figure 2. Phylogenetic tree based on the alignment of *nodC* gene sequences inferred using the maximum likelihood and general time reversible model. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the General time reversible model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. There were 444 positions in the final dataset.

KJ790183 Rhizobium leguminosarum ZY-305s (Ammopiptanthus sp.)
CP022669 Rhizobium leguminosarum sv. viciae BIHB 1217 (Pisum sativum)
FJ596038 Rhizobium leguminosarum sv. viciae USDA 2370^T (Vicia)
MT786700 3ZE (Pisum sativum)
MT786697 9ZE (Pisum sativum)
MT786696 11ZE (Pisum sativum)
MT786695 13ZE (Pisum sativum)
MT786693 15ZE (Pisum sativum)
JF759752 Rhizobium leguminosarum sv. viciae BIHB 1164 (Pisum sativum)
MT786694 14ZE (Pisum sativum)
HQ394248 Rhizobium pisi DSM 30132^T (Pisum sativum)
MG927338.1 Rhizobium ruizarguesonis UPM1133^T (Pisum sativum)
MT786699 4ZE (Pisum sativum)
JF759761 Rhizobium leguminosarum sv. viciae BIHB 1212 (Pisum sativum)
LC042284 Rhizobium KNa13 (Vicia villosa)
LC042283 Rhizobium KNa8 (Vicia villosa)
KX241906 Rhizobium leguminosarum sv. viciae li29 (Lathyrus intricatus)
GQ866899 Rhizobium leguminosarum V-1 (Vicia sativa)
MT786698 7ZE (Pisum sativum)
FJ596027 Rhizobium PEVF03 (Vicia faba)
NC 008381 Rhizobium leguminosarum sv. viciae 3841 (Pisum sativum)
KC608575 Rhizobium laguerreae FB206^T (Vicia faba)
JN580683 Rhizobium fabae CCBAU 33202^T (Vicia faba)
JQ585907 Rhizobium anhuiense CCBAU 23252^T (Vicia faba)
CP025505 Rhizobium leguminosarum sv. viciae UPM791 (Pisum sativum)
JF759757 Rhizobium leguminosarum sv. viciae BIHB 1192 (Pisum sativum)
MT786692 Z1 (Pisum sativum)
MF572089 Rhizobium leguminosarum sv. viciae AM11R (Vicia faba L. minor)
MT786701 2ZE (Pisum sativum)
KX241920 Rhizobium leguminosarum bv. viciae vd44 (Vicia disperma)
KX241938 Rhizobium leguminosarum bv. viciae vd14 (Vicia disperma)
JN649013 Rhizobium binae BLR235 (Lens culinaris)
JN649004 Rhizobium lentis BLR105 (Lens culinaris)
JN649011 Rhizobium bangladeshense BLR175^T (Lens culinaris)
FJ895269 Rhizobium leguminosarum bv. trifolii ATCC 14480 (Trifolium pratense)
KU664566 Rhizobium aegyptiacum 1010^T (Trifolium alexandrinum)
KF787869 Rhizobium vallis J15 (Phaesolus vulgaris)
KJ831243 Rhizobium sophorae CCBAU 03386^T (Sophora flavescens)
KJ921059 Rhizobium hidalgonense FH14^T (Phaesolus vulgaris)
HM441255 Rhizobium phaseoli ATCC 14482^T (Phaesolus vulgaris)
AF217268 Rhizobium etli sv. phaseoli CFN42^T (Phaesolus vulgaris)
sv. viciae
sv. trifolii
sv. phaesoli
86
99
97
99
96
96
94
79
81
98
70
100
100
89
100
90
0.10

Figure 3. Phenotypic features of *Rhizobium* isolates. Six *Rhizobium* isolates (2ZE, 9ZE, 13ZE, 14ZE, 15ZE, and Z1) were screened for their capabilities to use a. carbon (PM1–2) and b. nitrogen (PM3) sources and c. for their pH sensitivity [PM10 (wells A1–A12)]. Data are expressed in Arbitrary OmniLog Units (AOU) and represent the area under the kinetic curves after 96 h of incubation in the OmniLog instrument at 27 °C. The AOU values are reported in grayscale. For carbon and nitrogen sources (panel a and b), the scale ranges from white indicating the average AOU detected in the negative control [well A01] to black indicating the maximum AOU value. In panel c, the grayscale ranges from 10,000 AOU (white) to 100,000 AOU (black). For each condition, the reported AOUis the average obtained fromtwo replicates. d. Principal component analysis of PM data (PM1, PM2, PM3, PM9, and PM10) of *Rhizobium* isolates.



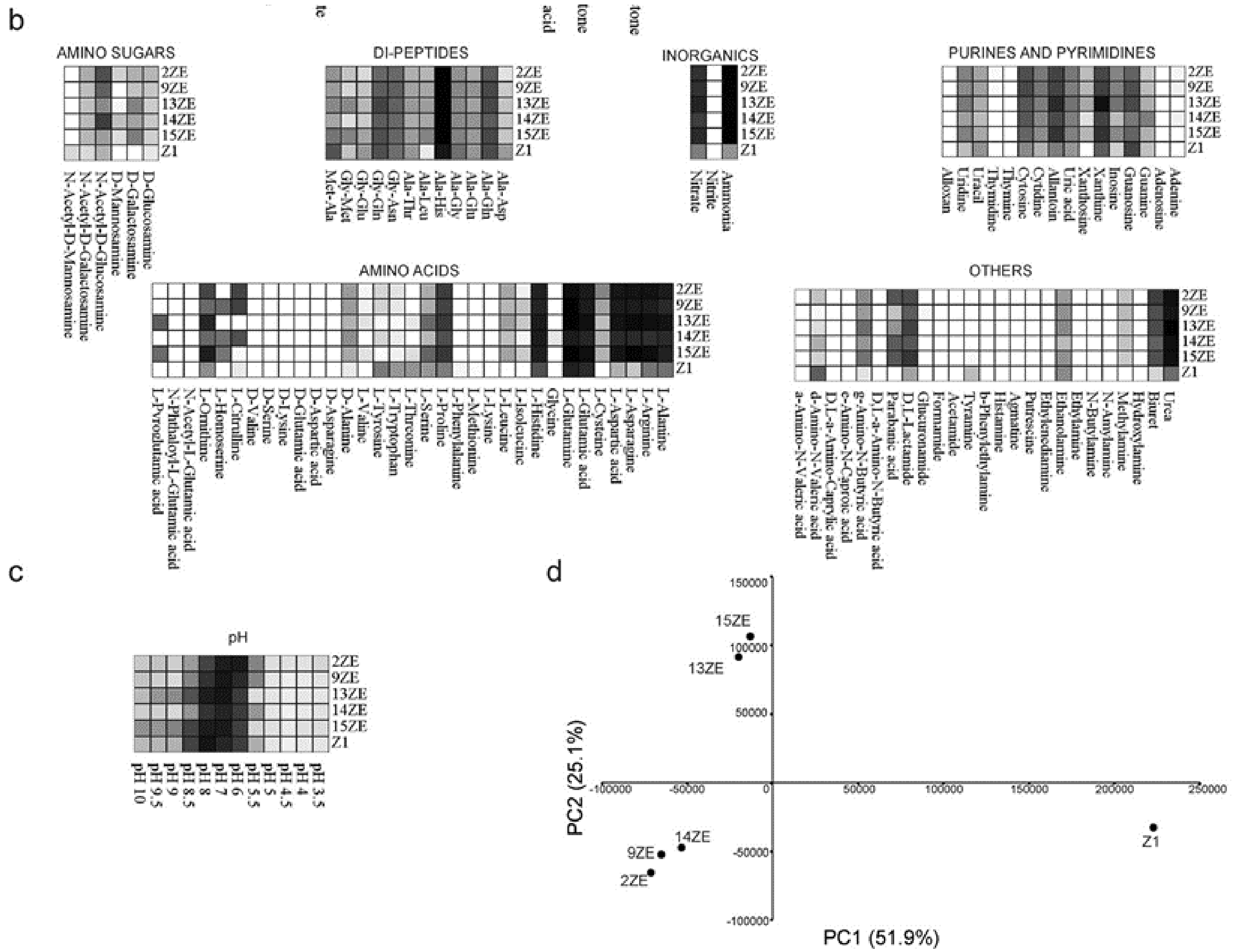
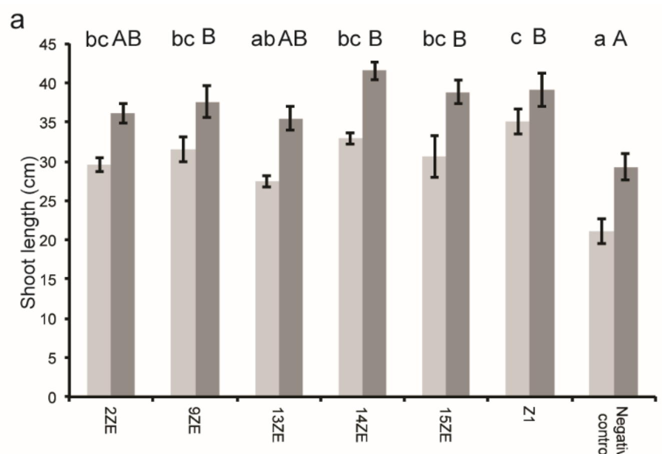
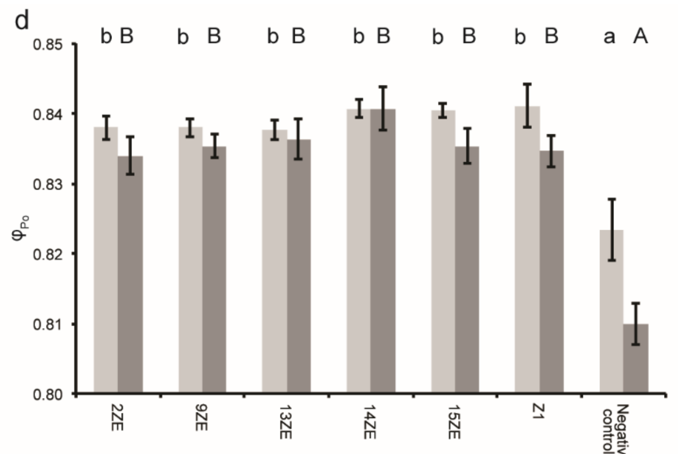
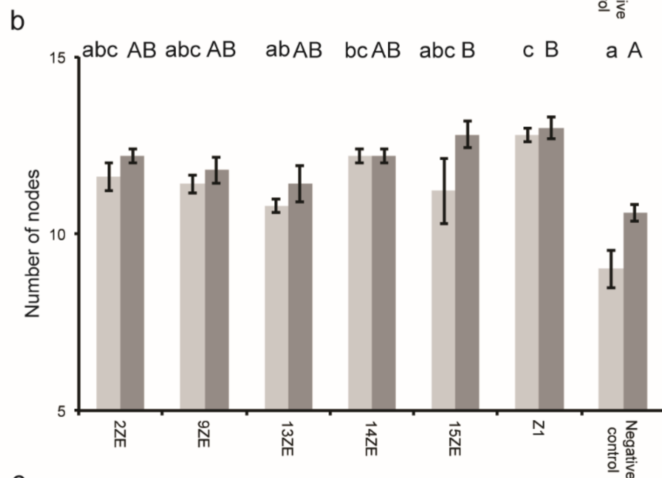
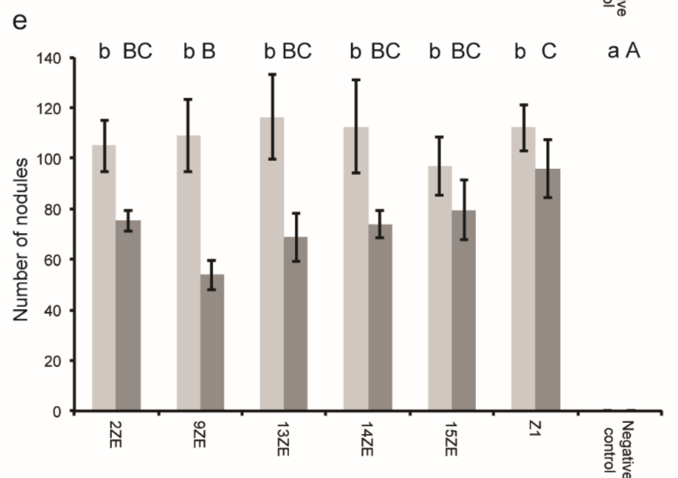

b
AMINO SUGARS
DI-PEPTIDES
INORGANICS
PURINES AND PYRIMIDINES
AMINO ACIDS
OTHERS
c
pH
d
PC1 (51.9%)
PC2 (25.1%)
15ZE
13ZE
14ZE
9ZE
2ZE
Z1

Figure 4. Plant parameters influenced by *Rhizobium* strains inoculation. a. Shoot length (cm), b. Number of nodes, c. Dry weight (g), d. φPo, the maximum quantum yield of primary photochemistry of a dark-adapted leaf e. Number of nodules of *Pisum sativum* plants inoculated or not with rhizobia strains. Light grey bars and dark grey bars indicate breeding lines 'DS 3637–2' and 'DS 3795–3', respectively. Average values calculated from five biological replicates are reported, with the standard error shown by error bars. Different letters (lower case letters for the breeding line 'DS 3637–2' and capital letters for 'DS 3795–3') indicate significant differences between strains ($p \leq 0.05$, Tukey HSD and Dunn test post hoc test for normally distributed and nonnormal data respectively).


a
bc AB
bc B
ab AB
bc B
bc B
c B
a A
Shoot length (cm)
2ZE
9ZE
13ZE
14ZE
15ZE
Z1
Negative control


d
b B
b B
b B
b B
b B
b B
a A
φ_{Po}
2ZE
9ZE
13ZE
14ZE
15ZE
Z1
Negative control


b
abc AB
abc AB
ab AB
bc AB
abc B
c B
a A
Number of nodes
2ZE
9ZE
13ZE
14ZE
15ZE
Z1
Negative control


e
b BC
b B
b BC
b BC
b BC
b C
a A
Number of nodules
2ZE
9ZE
13ZE
14ZE
15ZE
Z1
Negative control

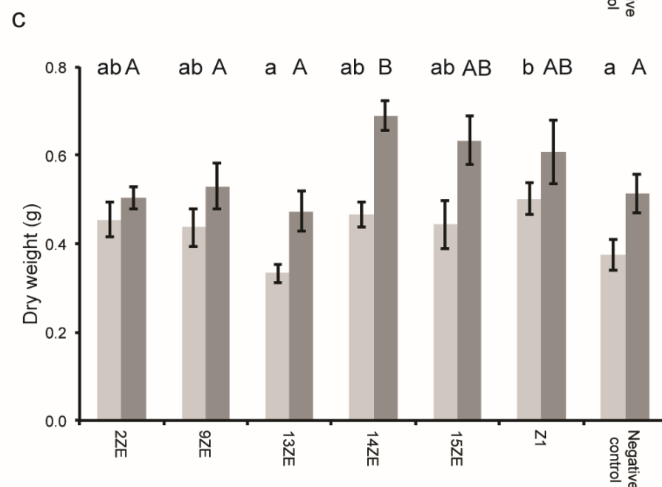

c
ab A
ab A
a A
ab B
ab AB
b AB
a A
Dry weight (g)
2ZE
9ZE
13ZE
14ZE
15ZE
Z1
Negative control

Figure 5. PCA of plants attributes inoculated with *Rhizobium* strains. Plants are indicated based on the inoculated strain. Shoot length (SL), number of nodes (ND), total plant dry weight (DW), photosynthetic yield (φPo), and nodule number (NDL)] were determined in a) *Pisum sativum* breeding line 'DS 3637–2' and b) *Pisum sativum* breeding line 'DS 3795–3' inoculated with *Rhizobium* isolates or not.

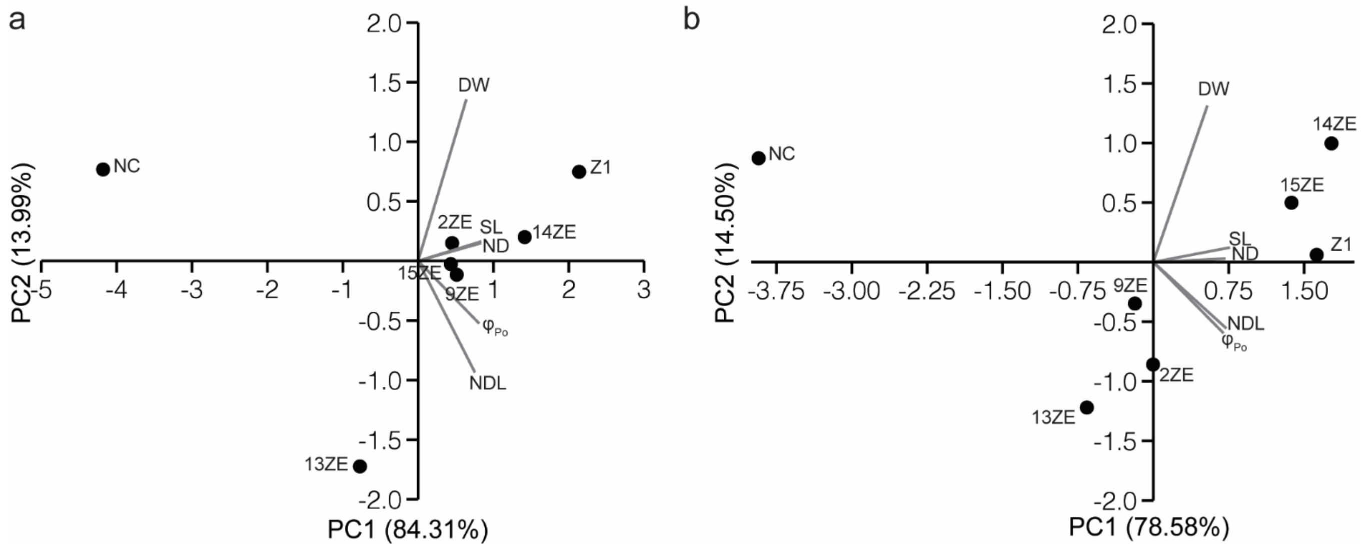
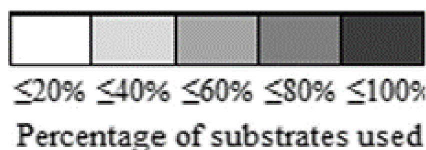


Table 1. Number of carbon and nitrogen sources used by *Rhizobium* strains. The substrates were divided into five and seven groups for carbon and nitrogen sources, respectively, based on their chemical classification. The total number of substrates tested for each class of compound is reported and the number of substrates used by each strain for each class of substrates is reported. The grayscale indicates the percentage of substrates utilized with each different class of compounds.

		Number of sources tested	2ZE	9ZE	13ZE	14ZE	15ZE	Z1
CARBON SOURCES	AMINOACIDS	*27*	6.0	7.0	10.0	7.0	12.0	8.0
	CARBOHYDRATES	*93*	49.0	52.0	48.0	50.0	51.0	53.0
	CARBOXYLIC ACIDS	*47*	4.0	4.0	8.0	8.0	12.0	13.0
	PEPTIDES	*5*	0.0	0.0	0.0	0.0	1.0	1.0
	OTHERS	*18*	1.0	1.0	2.0	2.0	2.0	0.0
	TOTAL	***190***	**60.0**	**64.0**	**67.0**	**66.0**	**77.0**	**75.0**
NITROGEN SOURCES	AMINOACIDS	*33*	16.0	17.0	16.0	16.0	18.0	13.0
	INORGANICS	*3*	2.0	2.0	2.0	2.0	2.0	2.0
	AMINOSUGARS	*6*	5.0	4.0	4.0	5.0	5.0	3.0
	PURINES	*9*	6.0	6.0	6.0	6.0	6.0	5.0
	PYRIMIDINES	*7*	4.0	4.0	4.0	4.0	4.0	4.0
	PEPTIDES	*12*	12.0	12.0	12.0	12.0	12.0	11.0
	OTHERS	*25*	8.0	7.0	8.0	8.0	8.0	7.0
	TOTAL	***95***	**53.0**	**52.0**	**52.0**	**53.0**	**55.0**	**45.0**

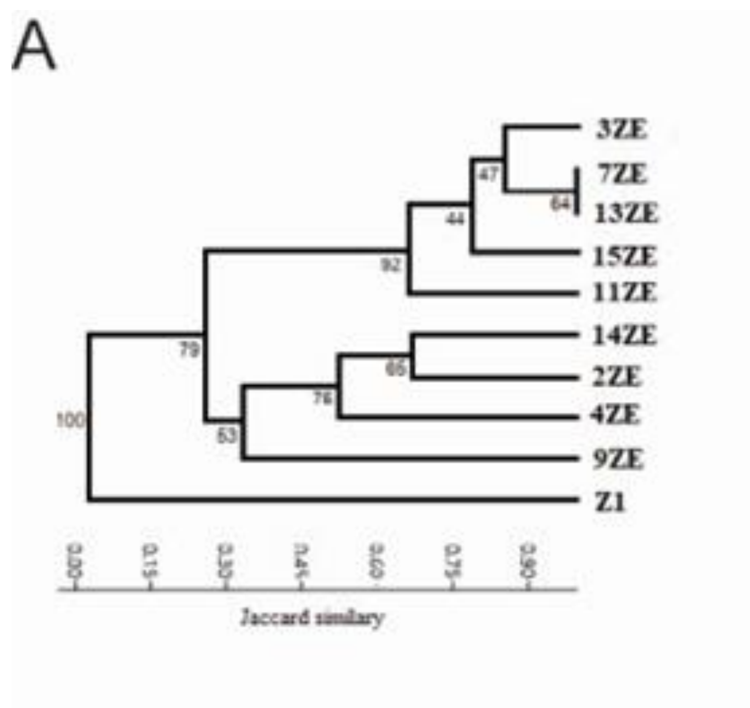


Supplementary material

Figure S1. Phylogenetic tree based on the alignment of 16S rDNA gene sequences. The evolutionary history was inferred by using the maximum likelihood method and Tamura 3-parameter model. The percentages of trees in which the associated taxa clustered together are shown next to branches (nodes below 70 not displayed). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. *Bradyrhizobium japonicum* LMG 6138^T (NR118981) 16S rDNA gene sequence was used as outgroup.

NR 118339 *Rhizobium leguminosarum* bv. *viciae* USDA 2370^T
JN558651 *Rhizobium laguerreae* FB206^T
MG904297 *Rhizobium ruizarguesonis* UMP1133^T
KJ921034 *Rhizobium hidalgonense* FH14^T
KJ921033 *Rhizobium acidisoli* FH13^T
MT775524 Z1
MT775522 14ZE
MT775519 9ZE
MT775517 4ZE
MT775515 2ZE
KF111868 *Rhizobium anhuiense* CCBAU 23252^T
88
KJ831229 *Rhizobium sophorae* CCBAU 03386^T
MT775516 3ZE
MT775518 7ZE
MT775520 11ZE
MT775521 13ZE
MT775523 15ZE
JN129381 *Rhizobium ecuadorense* CNPSo 671^T
DQ835306 *Rhizobium fabae* CCBAU 33202^T
AY509899 *Rhizobium pisi* DSM 30132^T
EF141340 *Rhizobium phaseoli* ATCC 14482^T
KC293513 *Rhizobium esperanzae* CNPSo 668^T
KJ438829 *Rhizobium chutanense* C5^T
U28916 *Rhizobium etli* CFN 42^T
JN648905 *Rhizobium lentis* BLR27^T
KJ831225 *Rhizobium sophoriradicis* CCBAU 03470^T
NR 137242 *Rhizobium binae* BLR195^T
NR 137399 *Rhizobium aegyptiacum* 1010^T
JN648931 *Rhizobium bangladeshense* BLR175^T
Y10170 *Rhizobium sullae* IS 123^T
AF364068 *Rhizobium indigoferae* CCBAU 71042^T
U86343 *Rhizobium gallicum* R602sp^T
95
U89817 *Rhizobium mongolense* 1844^T
85
FJ839677 *Rhizobium vallis* CCBAU 65647^T
EU867317 *Rhizobium endophyticum* CCGE2052^T
EU488750 *Rhizobium giardinii* H152^T
100
FJ154088 *Rhizobium sphaerophysae* CCNWGS0238^T
EU074168 *Neorhizobium alkalisoli* CCBAU 01393^T
GU128881 *Rhizobium vignae* CCBAU 05176^T
92
NR 113713 *Neorhizobium galegae* NBRC 14965^T
NR 118981 B*radyrhizobium japonicum* LMG 6138^T

Figure S2. The random amplified polymorphic DNA (RAPD) profile of *Rhizobium* isolates. A) Dendrogram based on UPGMA clustering of Jaccard similarity coefficient of combined AP5 and P7 RAPD fingerprints of rhizobial strains. Bootstrap values are given alongside the nodes (1000 replicates). Gel pictures of RAPD profiles with primer B) AP5 and C) P7.



B

C

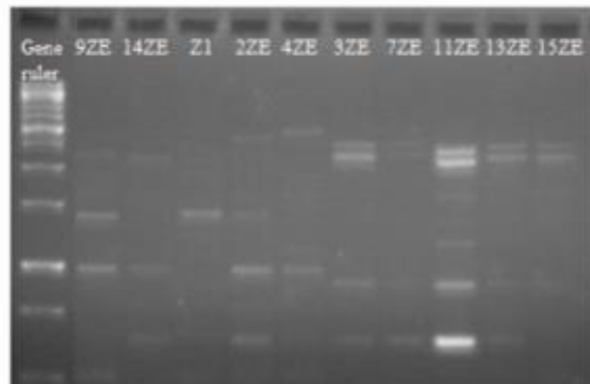


Figure S3. Phenotypic properties of *Rhizobium* isolates. Six *Rhizobium* isolates (2ZE, 9ZE, 13ZE, 14ZE, 15ZE and Z1) were screened for their capabilities to use a) carbon (PM1-2) and nitrogen (PM3) sources, b) osmolytes resistance (PM9) and for c) their pH sensitivity [PM10 (wells A1-A12)]. Heatmaps and clustering of *Rhizobium* isolates are reported. Data are expressed in Arbitrary OmniLog Units (AOU) and represents the area under the kinetic curves after 96 hours of incubation in the OmniLog instrument at 27°C. AOU are reported in gray scale. For carbon and nitrogen sources the scale ranges from white, indicating the average AOU detected in the negative control [well A01,] to black, indicating the maximum AOU value. In panel C the gray scale ranges from 10000 AOU (white) to 100000 AOU (black). For each condition, the reported AOU is the average obtained from two replicates. The two-way clustering of the strains and substrates was created by TMEV software using Pearson correlation coefficient based on average linkage method. (figure online on https://link.springer.com/article/10.1007%2Fs13199-021-00747-7#Sec13)

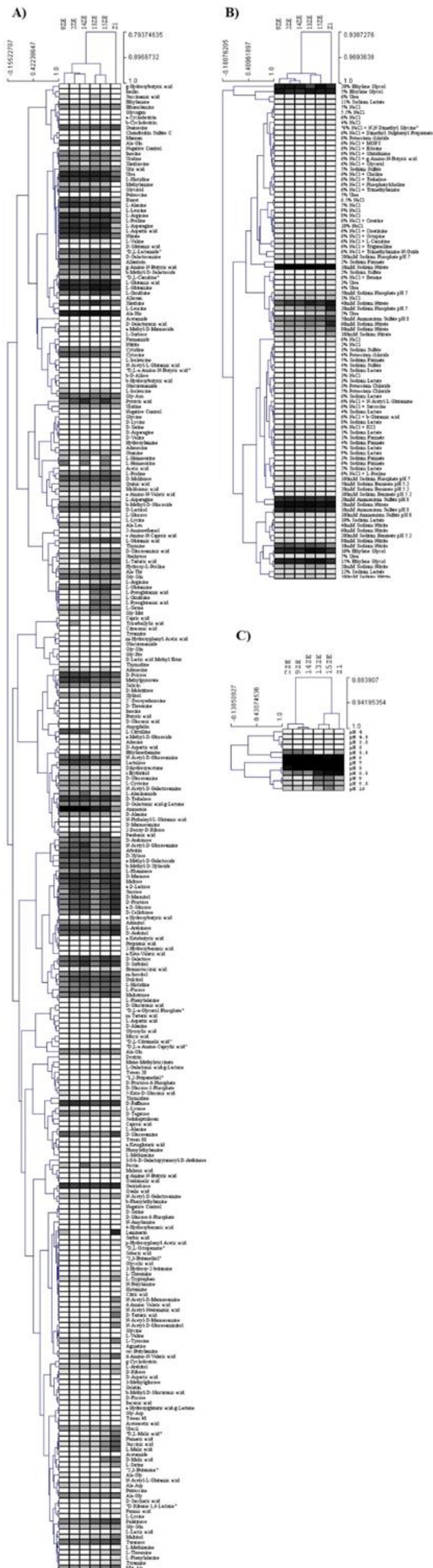
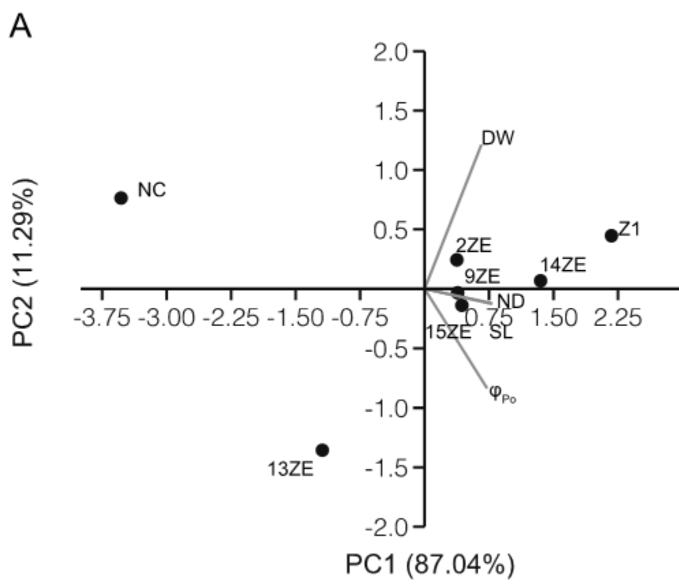

A)
B)
C)

Figure S4. PCA of attributes of plants inoculated with *Rhizobium* strains. Plants are indicated based on the inoculated strain. Shoot length (SL), number of nodes (ND), total plant dry weight (DW) and photosynthetic yield (φPo) were determined in a) *Pisum sativum* breeding line 'DS 3637-2' and b) *Pisum sativum* breeding line 'DS 3795-3' inoculated with *Rhizobium* isolates or not.


A
PC2 (11.29%)
PC1 (87.04%)
NC
DW
2ZE
9ZE
14ZE
Z1
ND
15ZE
SL
φPo
13ZE
2.0
1.5
1.0
0.5
-0.5
-1.0
-1.5
-2.0
-3.75 -3.00 -2.25 -1.50 -0.75
0.75 1.50 2.25

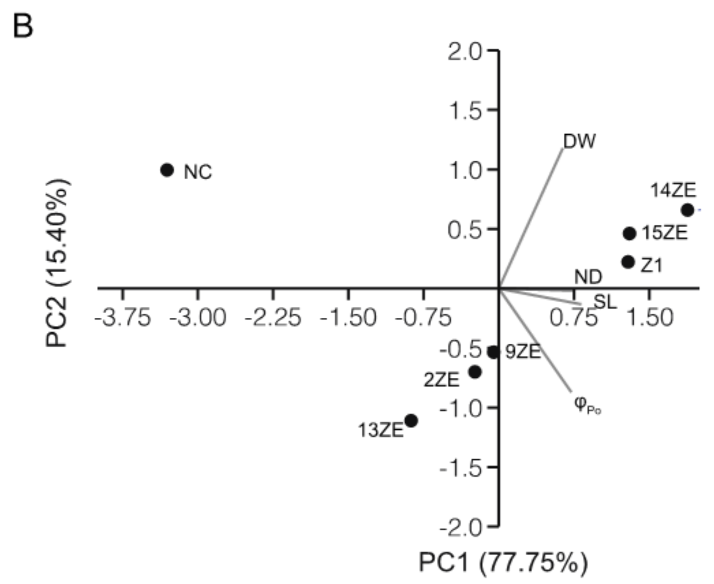

B
PC2 (15.40%)
PC1 (77.75%)
NC
DW
14ZE
15ZE
Z1
ND
SL
9ZE
2ZE
13ZE
φPo
2.0
1.5
1.0
0.5
-0.5
-1.0
-1.5
-2.0
-3.75 -3.00 -2.25 -1.50 -0.75
0.75 1.50

Table S1. Plant parameters significantly affected by *Rhizobium* isolates.

Strain	Breeding line	Shoot length	Number of nodes	Dry weight	φ_{Po} [1]	Number of nodules	Number of parameters[2]
2ZE	DS 3637-2	29.50±0.880*	11.60±0.400	0.45±0.039	0.84±0.002*	105.00±10.349*	4
	DS 3795-3	36.08±1.209	12.20±0.200	0.50±0.025	0.83±0.003*	75.40±4.226*	3
9ZE	DS 3637-2	31.50±1.541*	11.40±0.245	0.44±0.043	0.84±0.001*	109.00±14.265*	3
	DS 3795-3	37.62±2.095*	11.80±0.374	0.53±0.052	0.84±0.002*	53.80+±5.911*	3
13ZE	DS 3637-2	27.40±0.678*	10.80±0.200	0.33±0.021	0.84±0.001*	116.40±16.795*	2
	DS 3795-3	35.42±1.509*	11.40±0.510	0.47±0.045	0.84±0.003*	68.80±9.541*	2
14ZE	DS 3637-2	32.90±0.731*	12.20±0.200*	0.47±0.029	0.84±0.001*	112.60±18.419*	5
	DS 3795-3	41.52±1.100*	12.20±0.200	0.69±0.034*	0.84±0.003*	74.00±5.450*	6
15ZE	DS 3637-2	30.60±2.690*	11.20±0.917	0.44±0.054	0.84±0.001	96.80±11.495*	3
	DS 3795-3	38.84±1.552*	12.80±0.374*	0.63+±0.056	0.84±0.002	79.60±11.818*	4
Z1	DS 3637-2	35.10±1.616*	12.80±0.200*	0.50±0.037*	0.84±0.003*	112.20±9.2339*	7
	DS 3795-3	39.12±2.194*	13.00±0.316*	0.61±0.072	0.83±0.002	6.00±11.550*	4
Negative control	DS 3637-2	21.10±1.560	9.00±0.548	0.37±0.034	0.82±0.004	0.00±0.000	-
	DS 3795-3	29.30±1.655	10.60±0.245	0.51±0.044	0.81±0.003	0.00±0.000	-

[1] φPo: maximum quantum yield of primary photochemistry in the dark-adapted state.

2 Total number of plant parameters significantly affected by the inoculum of the *Rhizobium* strains respect to the negative control.

* Plant parameter significantly affected in the inoculated plants.

Chapter 6

Concluding remarks

The environmental and economic cost of the massive use of chemical nitrogen fertilizers in agriculture is a global concern. Besides, international emphasis on sustainable agriculture mandates that alternatives to nitrogen fertilizers must be urgently sought, increasing the attention on the potential role of BNF in supplying nitrogen for agriculture. It is well known that rhizobium inoculants are widely used in agriculture, providing one of the most environmentally sustainable and both cost-effective ways to boost legume performances. However, in the race to develop improved rhizobial inoculants, the screening for rhizobia with high nitrogen fixation efficiency cannot be the sole criterion for further field application. Indeed, the longstanding 'competition problem' detected in field harshly challenges the feasibility of the employed rhizobial bio-inoculants, limiting their positive effects on legume yields. Therefore, filling the gap about the main features that a priori makes rhizobial strains able to outcompete others indigenous rhizobia is of primary importance.

With this aim, we explored the competitiveness capabilities of *S. meliloti* for alfalfa nodule occupancy by using a panel of strains fluorescently labelled (Chapter 3). Our accurate experimental design revealed an extensive variety of *S. meliloti* strain response to the diverse competitive conditions, thus outlining a highly complex phenotype that strongly depends on the engaged competitor. Moreover, we have proven that predictive models, trained by using the k-mer based genome-wide association (GWAS) approach, can effectively predict the competition abilities of *S. meliloti* strains in the three observed competition patterns. Notably, this k-mer based GWAS approach was successfully applied for the first time to identify the genomic biomarkers controlling the competitiveness in *S. meliloti.* The uncovered genetic variants we reported, associated with remarkable competition capabilities, were mainly related to biosynthesis of cofactors, to transport and metabolism functions. Therefore, several functions contribute to ameliorate competitiveness indicating that many different bricks, increasing rhizobial versatility, pave the way for success in competition.

The phenotypic features linked to adaptation to different soil conditions are also determining for the formulation of effective inoculants. Therefore, in this thesis we also used a multiphasic approach for the evaluation of performances of *Rhizobium leguminosarum* strains isolated from pea plants, combining molecular characterization and Phenotype microarrays analysis with classic nodulation experiments (Chapter 5). Based on obtained results, two *R. leguminosarum* strains (*Rhizobium anhuiense* strain Z1 and *Rhizobium leguminosarum* bv. *viciae* 14ZE) were identified as the best candidate for inoculation of pea improving its growth.

Overall, the obtained results highlighted how the selection of rhizobia strains to be used as inoculants is a critical process that should paid more attention. For the first time, we have reported on the feasibility and reliability of using a k-mer-based GWAS approach for the genotype-to-phenotype association in a complex quantitative phenotype as competition for nodulation. Indeed, the genetic aspects associated with remarkable competition capabilities were tracked down providing useful information for the selection of both highly competitive and efficient rhizobia. Therefore, the k-mer based GWAS approach employed in this thesis may be helpful for large-scale screening of putative competitiveness capabilities among pairs of strains on the basis of the only genome sequences, to develop novel ameliorated inoculants. Moreover, using a classic characterization approach, we showed the importance of developing tailor suit inoculants for specific legume cultivars/lines.